5/17

CD-uitbreidingen: Video-CD

Inhoud

5/17.1

Het principe van Video-CD

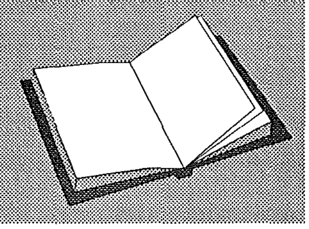
De voorgeschiedenis

Optimistische marktverkenningen
Marktonderzoekers zijn er al jaren van overtuigd dat er voor een CD-systeem, waarmee men films kan afspelen, een glansrijke toekomst is weggelegd. Volgens voorspellingen zou de markt voor een dergelijk medium in 1995 wereldwijd goed zijn voor een omzet van 16 miljoen exemplaren. Het is dus logisch dat de fabrikanten van consumer-elektronica al jaren lang bezig zijn met het bedenken van systemen, waarmee video-informatie op de kleine glanzende CD-schijfjes kan worden vast gelegd.

Voordelen
Het onderbrengen van video op CD heeft natuurlijk grote voordelen. Voor de consument geldt de kwaliteit. VHS-banden slijten en verliezen zelfs in opgeborgen toestand een deel van hun magnetische informatie. CD's zijn levenslang bruikbaar en slijten bij normaal gebruik niet. Ook voor de fabrikanten zou een dergelijk systeem voordelen hebben. VHS-banden worden, ondanks Macrovision, bij het leven gekopieerd waardoor een deel van de markt verloren gaat. CD's zijn nog niet te kopiëren, hoewel er natuurlijk waarschijnlijk binnen vijf jaar redelijk betaalbare systemen te koop zijn, waarmee CD's te beschrijven zijn. Een ander voordeel is echter dat de productie van een CD heel wat minder kost dan de productie van een VHS-kassette. Voor het bespelen van een VHS-band moet men ongeveer *f* 6,00 betalen. Het persen van een CD kost echter slechts ongeveer *f* 2,00!

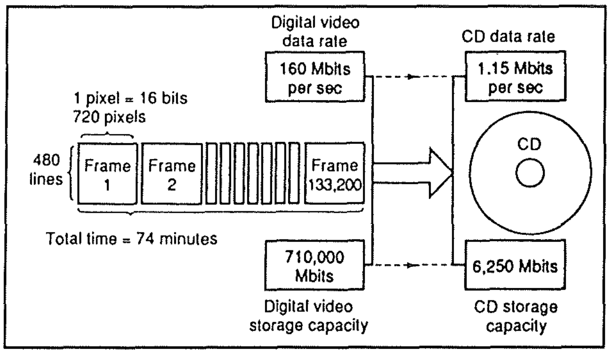
Een nieuwe standaard
In september 1993 werd door de vooraanstaande fabrikanten een nieuwe loot aan de uitgebreide stamboom van de CD-technologie gedefinieerd. In het zogenoemde "White Book" werden de specificaties vastgelegd van een systeem, waarmee het mogelijk is kwalitatief bevredigend video-informatie in de spiraal van een normale CD vast te leggen. Deze norm moest de verwarring die er op het gebied van "Video op CD" is ontstaan, door het op de markt verschijnen van talrijke systemen, in één klap oplossen. De

LEES OOK:
Hoofdstuk 5/8.1
Hoofdstuk 5/17.2
Hoofdstuk 5/17.3
Hoofdstuk 5/17.4



nieuwe internationale norm "Video-CD" was hiermee geboren.

Grote technologische problemen
Het ontwikkelen van een "Video op CD"-systeem is geen sinecure. Er zijn twee grote problemen: de voor video te lage opslagcapaciteit van een CD en de te lage doorvoersnelheid van gegevens. Beide problemen worden dramatisch geïllustreerd in figuur 5/17.1-1.



Figuur 5/17.1-1: Een CD-systeem voor video vereist een data-reductie met een factor 140.

Een gedigitaliseerde film met een speelduur van 74 minuten heeft een omvang van ongeveer 710 Gbit. Om het normale aantal beeldjes per seconde op het scherm te zetten moeten er per seconde ongeveer 160 Mbit data getransporteerd worden. Nu heeft een CD een opslagcapaciteit van slechts 6,25 Gbit en een doorvoersnelheid van 1,15 Mbit. Behalve de eigenlijke video-informatie moet men natuurlijk ook nog stereogeluid en tal van foutcompensatie- en instructiegegevens transporteren en opslaan. Men kan berekenen dat de omvang van de gegevensstroom met een factor 140 gereduceerd moet worden!
In de loop der jaren zijn er talrijke pogingen ondernomen om deze problemen op te lossen en een goed werkend en commercieel aanvaardbaar "Video op CD"-systeem op de markt te zetten. Een kort overzicht volgt in de volgende paragraafjes.

CD+G:
het allereerste grafisch CD-systeem
Weinig gebruikers van de Audio-CD weten dat dit systeem in principe de mogelijkheid biedt eenvoudige grafische gegevens op een scherm weer te geven. Philips en Sony, de opstellers van de allereerste CD-norm "Red Book", hebben in deze norm de mogelijkheid ingebouwd ongeveer 1.500 grafische statische schermen, met een resolutie van 288 bij 192 pixels, op één Audio-CD op te nemen. De digitale gegevens zitten verstopt in de woorden R, S, T, U, V en W van de zogenoemde "sub-code". De grafiekjes zien er ongeveer uit zoals een Teletext-pagina en kunnen 16 kleuren bevatten uit een palet van 4.096.
Ieder scherm kan in ongeveer 2,5 seconde uit de spiraal van de CD gelezen worden. CD's en CD-spelers die deze norm ondersteunen, moesten de codering "CD+G" krijgen. In de praktijk heeft dit systeem zich echter niet doorgezet, meer dan een paar CD+G-afspelers zijn niet op de markt verschenen en de software-industrie vond de geboden mogelijkheden niet interessant. JVC heeft het systeem wel toegepast voor Karaoke-titels, waarbij de grafische mogelijkheden werden ingezet voor het weergeven van songteksten op een monitor.

CDV: Compact Disk Video
In 1987 waagde Philips een poging om een analoog CD-videosysteem op de markt te brengen. CDV, afkorting van "Compact Disk Video", kon analoge vi-

17.1 Het principe van Video-CD

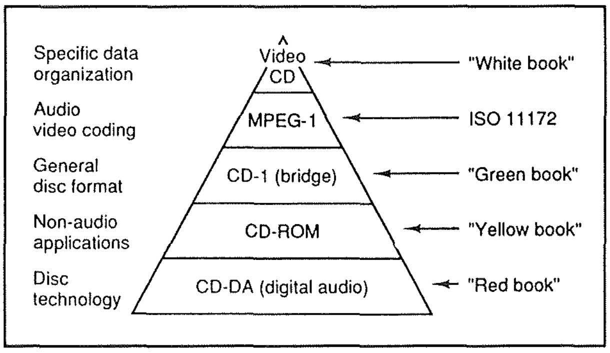
deogegevens en digitaal geluid op één CD opslaan. Vanwege de opslagproblemen kon een CD'tje echter slechts ongeveer 6 minuten video bevatten, waarbij er plaats voor ongeveer 20 minuten geluid vrij bleef. Ook dit systeem werd commercieel geen succes.

DVI: Digital Video Interactive

Het DVI-systeem werd in 1987 ontwikkeld door het Amerikaanse bedrijf RCA. Met DVI kan men één uur video op één CD'tje onderbrengen. Het systeem was in staat 800 kbit video samen te persen tot gemiddeld 5 kbit. Het systeem voldoet dus meer dan voldoende aan de vereiste compressie met een factor van 140. De toenmalige elektronica was echter niet in staat op een bevredigende manier de codeer- en decodeer-algoritmen op te lossen. Het comprimeren van één uurtje video kostte bijvoorbeeld meer dan een dag werk voor de allersnelste supercomputers! Het zal duidelijk zijn dat het prijskaartje dat hieraan vast hangt, de software-ontwikkelaars afschrikte. Bovendien slaagde RCA er niet in betaalbare decoders te ontwikkelen. Hiermee stierf DVI een vredige dood.

Software video

In 1985 werden de specificaties van CD-ROM door Philips en Sony vastgelegd in het zogenoemde "Yellow Book". Met dit systeem kan een CD door een PC opgevat worden als een normale schijf, waarop men met de door MS-DOS gedefinieerde bestandenstructuur digitale gegevens kan opslaan. Voor de overzichtelijkheid is in figuur 5/17.1-2 de geschiedenis van alle CD-standaarden als een piramyde voorgesteld. Diverse bedrijven ontwikkelden zuiver softwarematig werkende systemen, waarmee het mogelijk is video op een CD-ROM vast te leggen. Het bekendste systeem is ongetwijfeld het door MicroSoft ontwikkelde "Video for Windows", hoewel Intel met zijn "Indeo" een aardig partijtje mee blaast. Beide systemen hebben als nadeel dat de resultaten alleen op postzegelformaat in een venstertje van de monitor afgespeeld kunnen worden. Vergroot men de afmetingen, dan worden de beelden schokkerig opgebouwd en neemt de kwaliteit merkbaar af. Beide systemen zijn echter nog in ontwikkeling en in principe is het mogelijk dat zowel kwaliteit als afmetingen verbeterd worden tot er een bevredigend systeem ontstaat. Maar hoe dan ook, deze software-ontwikkelingen zullen nooit in staat zijn om echte films op een CD-ROM samen te persen.



Figuur 5/17.1-2: De piramyde van de ontwikkeling van de CD-standaarden.

CD-I: Compact Disk Interactive

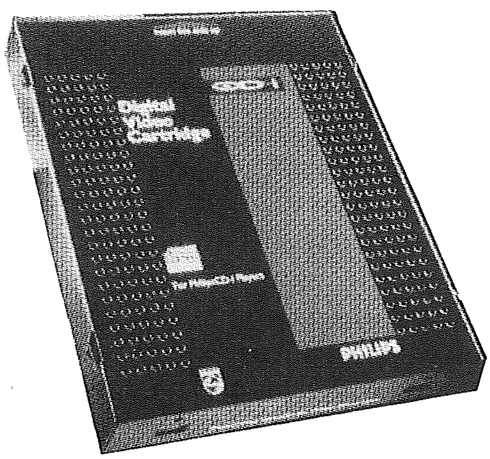
In 1988 werd de CD-boom uitgebreid met de "Green Book"-specificaties. Een systeem waaraan Philips de toekomst van haar consumer-divisie heeft opgehangen en dat van de CD-speler een interactief apparaat moet maken. Via CD-I kan men Multimedia-encyclopedieën raadplegen, cursussen volgen, talen leren, de diepzinnigste gedachten van popmuzikanten tot zich nemen en mooie dames in volledig eigen tempo beetje bij beetje ontkleden.

17.1 Het principe van Video-CD

De "Green Book"-specificaties ondersteunden video, zij het in een lage resolutie en met een maximaal aantal beeldjes van 15 per seconde. De bedoeling was dat deze video ondersteunend zou werken op het interactieve systeem en in een klein venstertje vertoond zou worden.
In 1988 gebeurde er echter iets, dat de "Green Book"-specificaties in één klap volledig ouderwets maakte. Een internationale groep van deskundigen stelde de MPEG-norm op, afkorting van "Moving Picture Experts Group". Deze norm beschrijft een zeer ingenieus compressiesysteem, waarmee het mogelijk is video-informatie te comprimeren tot opslag op CD met een redelijke kwaliteit mogelijk is. Met de MPEG-1 standaard kan men ongeveer 75 minuten video op één CD'tje vast leggen. De MPEG-standaarden zijn genormeerd in de internationale standaard ISO 11172.
Het was duidelijk dat MPEG ondersteuning verdiende in CD-I. Vandaar dat de oorspronkelijke "Green Book"-standaard werd aangepast, zodat er MPEG mee mogelijk werd.

CD-I FMV

Deze nieuwe standaard staat bekend als "CD-I FMV", afkorting van "Full Motion Video", oftewel volledig bewegende video. Er waren echter enige markttechnische problemen. Zoals bekend ging het in die jaren uiterst beroerd met het Philips-concern. De nieuwe leiding had de toekomst van Philips gekoppeld aan CD-I. Het kwam er dus op aan zo snel mogelijk CD-I spelers en (nog belangrijker) software op de markt te brengen. Het opstellen van de definitieve MPEG-norm wilde maar niet vlotten. Philips kon niet wachten en lanceerde CD-I voordat MPEG volledig uitgekristalliseerd was. Vandaar dat de allereerste generatie CD-I spelers zonder MPEG-ondersteuning op de markt kwamen. Wél werden de apparaten voorzien van een mogelijkheid om later een MPEG-decoder in te bouwen, de bekende DVC of "Digital Video Cartridge" die in figuur 5/17.1-3 is voorgesteld.



Figuur 5/17.1-3: De "Digital Video Cartridge", waarmee men CD-I spelers van de eerste generatie geschikt kan maken voor het afspelen van Video-CD's.

Een tweede probleem is dat de meeste CD-I titels die nu op de markt zijn geen ondersteuning bieden voor de definitieve MPEG-norm. Een probleem, waar in een van de volgende hoofdstukken dieper op wordt ingegaan.

Nimbus Video-CD: de verwarring stijgt ten top

Begin 1993, drie maanden nadat de definitieve MPEG-1 standaard was vast gelegd,

17.1 Het principe van Video-CD

verraste een vrijwel onbekend bedrijf, Nimbus Technology and Engineering, vriend en vijand door op de grote muziekbeurs MIDEM een volledig uitgewerkt Video-CD systeem te demonstreren. Het systeem ondersteunde MPEG-1 volledig en wat meer was, Nimbus beweerde dat de Video-CD'tjes op ***normale Audio-CD apparatuur*** die een digitale uitgang heeft afgespeeld kon worden. Op deze uitgang moest dan een losse MPEG-decoder aangesloten worden. Numbus demonstreerde op de beurs twee Video-CD's, die ieder meer dan anderhalf uur digitale video bevatten. De gecomprimeerde videogegevens waren in de normale audiosectoren van een Audio-CD geschreven. Volgens de stamvaders van de CD-technologie, Philips en Sony, overtrad het Nimbus-systeem echter de normen van het "Red Book". Dat is ook zo, want Nimbus gebruikte een niet standaard manier om de CD'tjes te beschrijven. De informatiedichtheid was namelijk groter, waarbij gehoopt werd dat de meeste moderne CD-spelers voldoende speelruimte hebben om die grotere informatiedichtheid toch zonder problemen van de spiraal uit te lezen.
Nimbus slaagde er niet in de software-industrie er van te overtuigen dat hun systeem goed was. Kopschuw gemaakt door de grote jongens, durfde men het niet aan het Nimbus-systeem softwarematig te ondersteunen. Dat betekende het einde van dit spectaculaire systeem.

Video-CD: de officiële norm

Op 29 juni 1993 werd de officiële Video-CD norm ten doop gehouden. Het waren niet de minste bedrijven die zich aaneen sloten: natuurlijk Philips en Sony, maar ook JVC, Matsushita met hun consumermerken Panasonics en Technics sloten een overeenkomst die alle verwarring rond het thema "Video op CD" moest uitsluiten. De officiële Video-CD norm werd vastgelegd in het "White Book" en ondersteunt niet alleen MPEG-1, maar zowel de Amerikaanse NTSC-standaard als het Europese PAL. Video-CD is in staat 74 minuten video op een normaal CD'tje vast te leggen en biedt daarnaast de mogelijkheid twee muziekkanalen in CD-kwaliteit aan te leveren.

"White Book" version 1.1

In januari 1994 werd de bestaande "White Book"-norm nog iets uitgebreid. Dit heeft echter geen consequenties voor het afspelen van Video-CD's. Door die uitbreidingen wordt het mogelijk ook "Still Video", stilstaande beelden, op een Video-CD op te nemen.
Deze beelden kunnen een resolutie hebben van 704 bij 480 pixels en er gaan er ongeveer 10.000 op één CD'tje! Bovendien ondersteunt versie 1.1 meerdere geluidsstructuren en instructie-code's zodat de "Still Video's" van kommentaar voorzien kunnen worden. De instructie-code's laten een interactieve toegang tot de Video-CD toe. Deze uitbreiding geeft Video-CD erg veel mogelijkheden op het gebied van onderwijs, instructie en voorraadbeheer en -herkenning.
Een verdere uitbreiding is het ondersteunen van "Dolby Surround Sound", waardoor het mogelijk wordt de videofilms, net als in de bioscoop, te voorzien van vier geluidskanalen. Hiervoor heeft de gebruiker natuurlijk wel weer een nieuw apparaat nodig dat in de bestaande geluidsinstallatie geïntegreerd moet worden en het extra geluid naar twee achter de toeschouwer opgestelde boxen voert. De huiskamer wordt langzaam maar zeker een echt controlecentrum!

De eigenschappen van Video-CD

Specificaties
MPEG-1 en dus ook Video-CD, is ontwikkeld voor de consumentenmarkt. Dat betekent ook consumenten kwaliteit. MPEG-1 levert beelden die een kwaliteit hebben die ***"vergelijkbaar is met deze van een goede VHS-recorder"***, althans volgens de ontwerpers. Mensen, die Video-CD's hebben bewonderd, menen echter dat de gemiddelde kwaliteit iets lager ligt. Het ingewikkelde compressie-algoritme dat wordt toegepast gooit heel wat beeldinformatie als zogenaamd overbodig weg. Bij sommige passages in een film blijkt dit zich te wreken op de beeldkwaliteit.
In het kort samengevat werkt Video-CD met onderstaande specificaties:

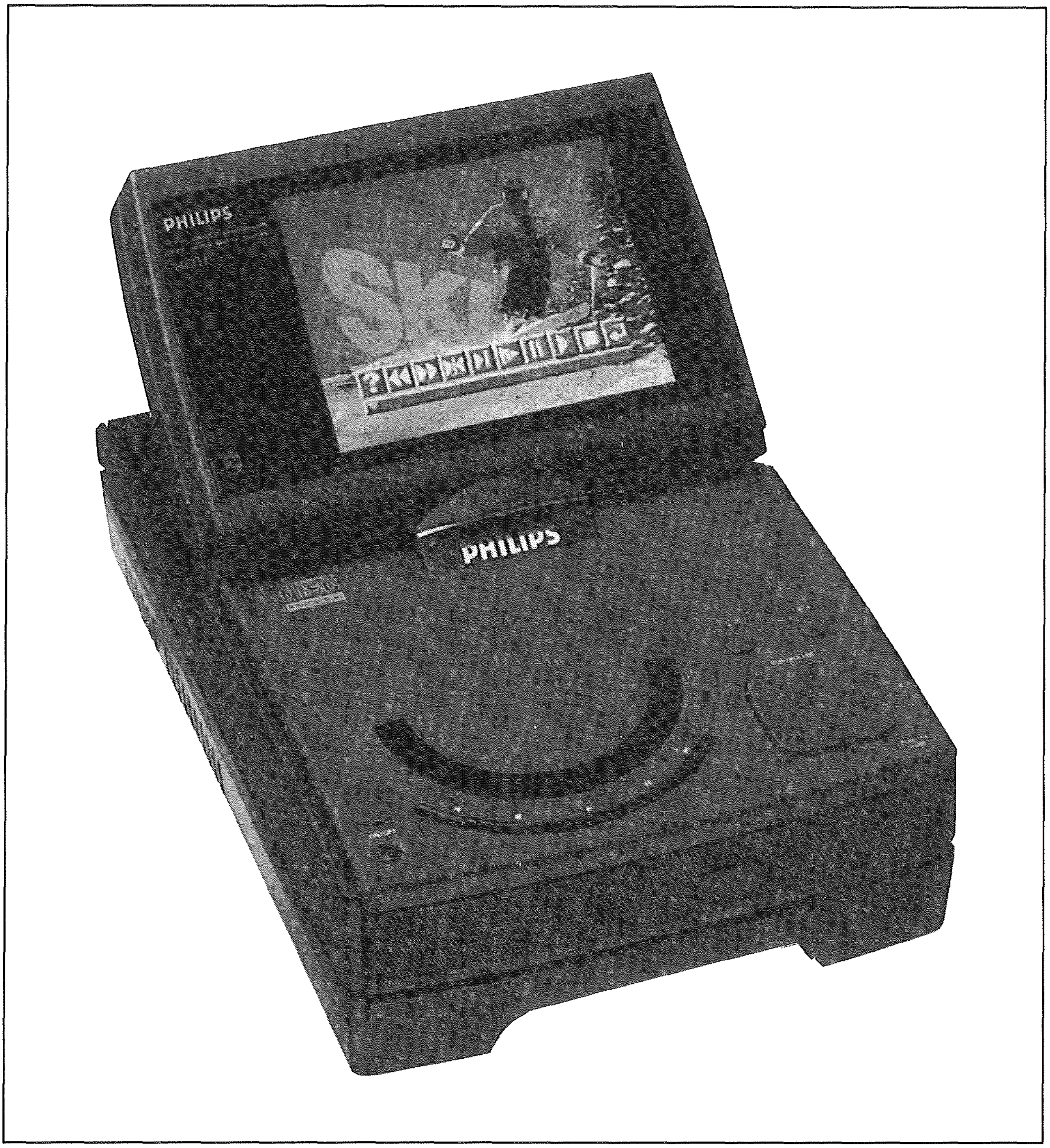
- **Het video gedeelte**
  - coderingssysteem:
    MPEG-1
  - resolutie:
    352 bij 240 voor NTSC
    352 bij 288 voor PAL
  - aantal beelden per seconde:
    30 voor NTSC
    25 voor PAL
  - overdrachtsnelheid:
    1,151 Mbit/s

- **Het audio gedeelte**
  - coderingssysteem:
    MPEG-1, vergelijkbaar met PASC
  - aantal samples per seconde:
    44,1 ks/s
  - overdrachtsnelheid:
    224 kbit/s
  - modes:
    twee kanaals mono of stereo

Hardware ondersteuning
Het was er de grote concerns natuurlijk alles aan gelegen om de Video-CD's op zo veel mogelijk reeds bestaande apparatuur afspeelbaar te maken. Dat is aardig gelukt:

- CD-I spelers
  Video-CD's zijn zonder problemen op CD-I spelers af te spelen, zij het dat de oude apparaten moeten voorzien worden van de Full Video Motio uitbreiding, de DVC-kassette. Vanwege de niet volledige ondersteuning van MPEG-1 door de "Green Book"-standaard zijn er enige kleine probleempjes, zie later.
- Computers
  PC's die voorzien zijn van een CD-ROM speler, die voldoet aan de XA-standaard Form-2 kunnen Video-CD's verwerken. Natuurlijk moet de PC dan worden uitgebreid met een MPEG-kaart, die echter steeds goedkoper wordt. Het is op dit moment niet te verwachten dat er softwarematige MPEG-1 decoders ontwikkeld zullen worden. Zelfs een 100 MHz Pentium processsor kan het in real time decoderen van de MPEG-1 datastroom niet aan! De MPEG-kaart moet via een kabeltje verbonden worden met de VESA-uitbreidingsconnector op de bestaande VGA-kaart. Diverse fabrikanten van grafische kaarten zijn bezig gecombineerde kaarten te ontwikkelen, waarbij de MPEG-1 decoder geïntegreerd is in de SUPER-VGA elektronica. Deze kaarten zullen ongetwijfeld echter alleen met een PCI-connector op de markt komen, zodat men een vrij modern moederbord in de PC moet hebben om deze kaarten te kunnen toepassen.
- Video-CD spelers
  Alle grote concerns zijn bezig speciale Video-CD spelers te ontwikkelen.

17.1 Het principe van Video-CD



Figuur 5/17.1-4: Een draagbare Video-CD speler, zoals die binnenkort in de winkels zal liggen.

Deze apparaten zijn in staat Video-CD's en Photo-CD's af te spelen, maar ***geen CD-I's***. Uiteraard zijn deze spelers standaard voorzien van MPEG-1 decoders. Ook draagbare Video-CD spelers zullen binnenkort te koop zijn. Deze apparaten zijn voorzien van een kleuren-LCD scherm. In figuur 5/17.1-4 wordt een van deze modellen, nu nog een prototype van Philips, voorgesteld.

17.1 Het principe van Video-CD

- Commodore's CD32
  De Amiga-computers met CD32 ondersteuning kunnen zonder problemen Video-CD's verwerken.
- Atari Jaguar
  De spelletjesconsole Jaguar van Atari is in staat Video-CD's te verwerken. Het combineren van interactieve Video-CD volgens de norm 1.1 met spelletjes zal een revolutie teweeg brengen op het gebied van levensechte spelletjes. De Video-CD kan zorgen voor levensechte achtergronden en karakters, het interactieve systeem zal toelaten dat de gebruiker de evolutie van de film of het spel zélf kan besturen.
- Sega
  Ook het spelletjesbedrijf Sega is klaar voor de komst van interactieve Video-CD. Hun spelletjescomputer Mega-CD is voorbereid op het opnemen van een MPEG-1 decoder.
- Software ondersteuning
  Op dit moment wordt Video-CD ondersteunt door de voornaamste leveranciers van films. Grote namen zoals Paramount, Polygram (Philips), Warner Bross, MGM en BMG zijn op dit moment bezig hun allernieuwste films onder de vorm van Video-CD's op de markt te brengen.

Toch enige kleine probleempjes ...

Niets lijkt een glorieuze opmars van Video-CD in de weg te staan. Men verwacht dat dit systeem binnen vijf jaar even ingeburgerd zal zijn als Audio-CD en dat VHS zijn langste tijd gehad heeft. Toch zijn er enige probleempjes, die alles te maken hebben met de stormachtige technische ontwikkelingen en het ongeduld van de grote bedrijven.

Het eerste probleem is de maximale speelduur van ongeveer 74 minuten. Het gevolg is dat een normale film niet op één Video-CD past, maar dat men altijd twee schijfjes nodig heeft. Op dit moment is nog niet te voorzien hoe de consument hierop zal reageren. Optimisten beweren dat mensen ook in de bioscoop gewend zijn een film op het spannendste moment te zien onderbreken voor een pauze. Hoe dan ook, diverse bedrijven hebben aangekondigd Video-CD wisselaars op de markt te brengen, die de twee schijfjes van één film opnemen en er volledig automatisch voor zorgen dat het beëindigen van het eerste schijfje vrijwel zonder onderbreking wordt gevolgd door de start van het tweede schijfje.

Elk jaar een nieuw systeem?

Een tweede probleem is dat de consument ervan overtuigd moet worden dat de aankoop van een CD-I of Video-CD speler een toekomstgerichte investering is. Dat zal niet erg lukken, nu Philips en Sony in november 1994 hebben aangekondigd een nieuwe internationale werkgroep op te willen richten ***"met als doel het vaststellen van een internationale norm voor de digitale opslag van video-informatie op optische dragers, waarvan de opslagcapaciteit in niet geringe mate deze van Video-CD zal overtreffen"*** (einde citaat). Hetgeen dus betekent dat de industrie, op het moment dat Video-CD's nog niet eens in de winkels liggen, alweer op zoek is naar een beter systeem. Hetgeen bovendien betekent dat iedereen die nu een Video-CD speler koopt, misschien binnen twee jaar met een volstrekt verouderd apparaat opgescheept zit.

De Video-CD/CD-I incompatibiliteit

Het derde probleem dat vele kandidaatkopers kan afschrikken is de CD-I/Video-CD incompatibiliteit. Zoals reeds beschre-

17.1 Het principe van Video-CD

ven heeft Philips met het uitbrengen van CD-I niet gewacht totdat MPEG-1 volledig gedefinieerd was. In augustus van 1993 werd een grote deal tussen Philips en Paramount bekend gemaakt, waardoor Philips het recht kreeg ongeveer vijftig kaskrakers van deze grote filmproducent op CD-I te zetten.

Het concern begon daar voortvarend aan, met als logisch gevolg dat er nu in feite twee verschillende videosystemen op CD bestaan. De Paramount-films voldoen niet aan de "White Book"-standaard van Video-CD, maar wel aan de "Green Book"-standaard van CD-I. Dat betekent dat deze films ***niet op een CD-ROM speler die werkt volgens de XA-standaard Form-2 afgespeeld kunnen worden***! Ook de toekomstige Video-CD spelers kunnen geen CD-I video's afspelen! Anderzijds is het zo dat CD-I spelers wél in staat zijn beide formaten te lezen.

Maar er is meer. De allereerste generatie CD-I spelers zijn voorzien van videostandaarden die iets afwijken van deze die door MPEG-1 onderschreven worden. De uitgebreide "Green Book" standaard specificeert bijvoorbeeld een resolutie van 384 pixels per beeldlijn, terwijl het "White Book" 352 pixels voorschrijft.

De kloksnelheid van een CD-I speler bedraagt 15 MHz, terwijl de gestandaardiseerde kloksnelheid van Video-CD 13,5 MHz bedraagt. Een en ander heeft tot gevolg dat Video-CD's, afgespeeld op een eerste generatie CD-I speler met DVC-uitbreiding, een smalle zwarte band aan weerszijden van de film op het scherm zetten en het beeld iets vervormd is. Philips heeft echter aangekondigd dat de nieuwste versie van haar DVC-unit wordt voorzien van een automatische schakelaar, die kan ontdekken of er een CD-I dan wel een Video-CD wordt afgespeeld. De klokfrequentie van het systeem wordt automatisch aangepast en het beeld iets uitgerokken, waardoor de problemen verholpen zullen zijn.

Een uniek Video-CD logo

Om de consumenten duidelijk te maken dat een film volgens de MPEG-1 standaard is opgenomen op een Video-CD die beschreven is volgens de "White Book"-standaard, hebben de fabrikanten van software afgesproken een uniek logo in te voeren voor dergelijke CD's. Dit logo is voorgesteld in figuur 5/17.1-5.



Figuur 5/17.1-5: Het unieke logo, waaraan men "echte" Video-CD's kan herkennen. Deze schijfjes kunnen afgespeeld worden op CD-I spelers, XA Form-2 CD-ROM drives en natuurlijk de speciale Video-CD spelers.

Video-informatie die volgens de "Green Book"-standaard op CD-I schijfjes staat, wordt op de hoes aangekondigd met de kreten "DIGITAL VIDEO ON CD-I" of "CD-I DIGITAL VIDEO".

Het NTSC/PAL-dilemma

Een volgend probleem is dat Video-CD schijfjes zowel bruikbaar zijn voor de Amerikaanse NTSC- als voor de Europese PAL-norm. Dat probleem is nieuw voor de video-producenten, omdat Amerikaanse voorbespeelde VHS-banden niet bruikbaar zijn in Europa. Nu kan men zich de vraag stellen waar het probleem zit. Dat is zuiver commercieel en heeft te maken met de release-data van nieuwe (Amerikaanse) films in Europa. Amerikaanse filmproducenten hebben de neiging hun films eerst in Amerika uit te brengen en nadien, soms met maanden vertraging in Europa.
Een en ander kan echter betekenen dat er in Amerika al Video-CD's op de markt verschijnen van een film die in Europa nog niet in de bioscopen is vertoond. Het is duidelijk dat de zwarte handel daar zou op inspelen, hetgeen de uitbaters van Europese bioscopen natuurlijk niet op prijs zouden stellen. Bovendien kunnen er nog problemen ontstaan, doordat de Amerikaanse normen en zeden nu eenmaal veel strenger zijn dan de Europese. Films, die in Europa zonder problemen vertoond worden, moeten in Amerika vaak bevrijd worden van "al te nadrukkelijk seksuele, erotische of godslasterlijke scènes". Het is niet de bedoeling dat dergelijke aanstootgevende scènes de Grote Oceaan oversteken en de onbevlekte Amerikaanse geest bederven!
Om al deze problemen op te lossen heeft men er voor gekozen Video-CD's te voorzien van blokkeercodes.
Deze zorgen ervoor dat een Amerikaanse Video-CD speler alleen Video-CD's kan afspelen die een NTSC-code hebben en Europese apparatuur alleen schijfjes kan verwerken die voorzien zijn van een PAL-code.

Het besturingssysteem van Video-CD spelers

Inleiding

In Video-CD en CD-I spelers is uiteraard een flink brok firmware aanwezig. De videogegevens zijn in een nogal vreemd formaat in de spiraal ondergebracht, waardoor in principe eenvoudige functies, zoals stilstaand beeld, niet zonder flink wat software uit te voeren zijn.
Het besturingssysteem heet "Full Motion CD-I" en dit zorgt ervoor dat de kijker/luisteraar op een interactieve manier gegevens van de Video-CD kan halen.

Audio en video controle

Een van de voornaamste doelstellingen van het besturingssysteem is de videostroom te koppelen aan een bepaalde audiostroom.
Het is ook mogelijk de MPEG-1 decodering van audio en video afzonderlijk te sturen. Voor beide systemen bestaat de mogelijkheid gegevens rechtstreeks van de CD naar het scherm te sturen, maar ook de gegevens eerst in het RAM-geheugen van de speler onder te brengen en vandaar af te spelen.
Deze twee functies worden respectievelijk "Disc Playback" en "Memory Playback" genoemd.

Besturingsfuncties

De "Full Motion CD-I" software biedt in totaal vijf functies waarmee de manier waarop de gegevens op het scherm verschijnen kan worden bestuurd.

- Play Forward
  De video-informatie wordt met de normale snelheid van 25 beeldjes per seconde naar het scherm gestuurd.

17.1 Het principe van Video-CD

- Slow Motion Forward
  De besturingssoftware biedt de mogelijkheid de beeldjes vertraagd naar het scherm te sturen, waarbij men kan kiezen uit een van zeven snelheden. Om flikkerende beelden te voorkomen wordt ieder beeldje langere tijd dan de normale 20 ms vertoond.
- Freeze Mode
  In deze modus wordt één beeldje op het scherm bevroren en wordt de decodering van de overige MPEG-1 beelden gestopt. De datastroom van de Video-CD gaat echter ongestoord door. Bij het omschakelen naar de "Play Foreward" mode wordt ingehaakt op de beelden die op dit moment van de CD worden gelezen. Het is echter niet mogelijk gelijk welk beeld te bevriezen. Alleen I-beelden (zie hoofdstuk 5/17.2), waarvan alle video-informatie gecodeerd is, kunnen stilstaand op het scherm verschijnen.
- Single Forward Mode
  Bij deze modus worden alle beelden een na een op het scherm vertoond. Het is mogelijk de duur van ieder schermbeeldje te programmeren.
- Scan Mode
  Nu worden alleen alle I-frames vertoond en wel met de snelheid waarmee zij in de video datastroom voorkomen.

17.1 Het principe van Video-CD

5/17.2

De MPEG-1 compressie van beeld en geluid

Inleiding

Compressie noodzakelijk

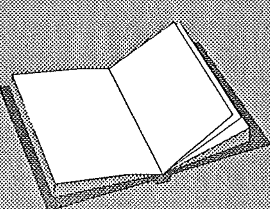
Toen de internationale communicatie-industrie in het begin van de jaren '80 begon te filosoferen over digitale communicatie werd al snel duidelijk dat het nooit zou lukken foto's, ander beeldmateriaal, video en geluid op een economische manier ongecomprimeerd te versturen. Immers, een gedigitaliseerde foto met een resolutie van 1.024 bij 1.024 pixels en met een kleurendiepte van 24 bit, een minimale vereiste voor professioneel kleurendrukwerk, neemt ongeveer 3 MB aan opslagruimte in beslag. Zou men een dergelijke foto via een 9.600 bps modem versturen, dan zou dit ongeveer 55 minuten verzendtijd kosten. Nu waren er wel een heleboel compressietechnieken ontwikkeld, maar de internationale industrie was er (gelukkig) van overtuigd dat er een wereldstandaard voor de compressie van digitale beelden en geluid moest opgesteld worden. Die standaard zou flexibel moeten zijn, zodat een en dezelfde hardware door software geprogrammeerd kon worden voor verschillende soorten compressie-algoritmen. Op deze manier zou de internationale standaard zonder grote problemen aangepast kunnen worden aan verschillende toepassingen, die ieder hun eigen eisen stellen.

JPEG

Op initiatief van twee internationale normeringsinstituten, de ISO (International Standardisation Organisation) en de IEC (International Electrotechnical Committee) werd in 1988 een internationale werkgroep opgericht met als naam "Joint Photographic Expert Group". Deze werkgroep ging door het leven onder de naam JPEG en was samengesteld uit internationaal toonaangevende experts op het gebied van compressietechnieken en digitale beeldverwerking. In de werkgroep waren alle belangrijke internationale industrieën vertegenwoordigd uit de gebieden consumentenelektronica, automatisering, datacommunicatie en IC-fabricage. De werkgroep ontwikkelde op korte termijn een beeldcompressietechniek, die geschikt is voor een breed scala van toepassingen met verschillende kwaliteiten, resoluties, compressieniveaus en kleurmodellen. Deze techniek kreeg de naam

LEES OOK:

Hoofdstuk 5/8.1
Hoofdstuk 5/17.1
Hoofdstuk 5/17.3



"JPEG" en werd beschreven in de ISO/IEC-norm 10.918. De werkgroep moest de techniek niet van de bodem af ontwikkelen. Een heleboel voorbereidend werk was reeds verricht door de "Philips Communicatie Industrie", die een beeldcompressietechniek had ontwikkeld die toegepast kon worden bij ISDN-netwerken die werken met een snelheid van 64 kb/s. Deze techniek maakt deel uit van de overdrachtstandaard H.261. De JPEG-norm sluit hierbij naadloos aan, zodat deze techniek voor beeldcompressie een grote toekomst tegemoet gaat, onder meer voor foto-archivering, telefonische beeldverzending, kleurfaxen, digitale fotocamera's, desktop publishing, digitale kopieerapparatuur en beeldtelefoons.

Statische beelden

JPEG werd in eerste instantie ontwikkeld voor de compressie van statische beelden, zoals krantenfoto's. Natuurlijk dacht men er wel aan deze techniek ook toe te passen voor het comprimeren van dynamische beelden, dus video. In de begintijd van deze techniek waren de noodzakelijke processoren echter niet snel genoeg om beelden in real time te comprimeren en nadien weer te decomprimeren. Het verwerken van een hoge resolutie afbeelding met 24 bit kleurendiepte kostte diverse minuten. Door de ontwikkelingen op het gebied van snelle processoren zou het op dit moment wél mogelijk zijn met JPEG video met 25 beelden per seconde real time te verwerken. Men zou zich dus de vraag kunnen stellen waarom het noodzakelijk was voor video een nieuwe techniek te ontwikkelen. Dat heeft alles te maken met de snelle evolutie in de consumentenelektronica. Het werd duidelijk dat de markt voor afspelers van Audio-CD's snel verzadigd zou geraken en dat het noodzakelijk was iets nieuws te verzinnen om de consument weer tot investeren te bewegen. Dat nieuwe zou Video-CD moeten worden, een systeem waarbij speelfilms op de bekende 12 cm CD's gedistribueerd zouden kunnen worden. Zelfs de snelste real time JPEG-compressie is hier niet geschikt voor. De CD-technologie werkt uit de aard der zaak serieel. De gegevens worden op bit-niveau een na een in de spiraal van de CD geschreven. Als men dus video op een CD wil opnemen, dan moeten de gecomprimeerde gegevens van alle beeldjes én van het stereogeluid omgezet worden in een seriële datastroom. Als er 25 beeldjes per seconde met JPEG gecomprimeerd worden, is de daaruit voortvloeiende seriële datastroom veel te groot om door een CD verwerkt te worden. Bovendien moet er natuurlijk ook nog eens gecomprimeerd digitaal geluid verzonden worden!

Een CD heeft echter een maximale seriële datasnelheid van 1,411 Mbit/s, vastgelegd door de specificaties die door Philips en Sony in het zogenoemde "Red Book" voor Audio-CD werden gedefinieerd. Deze maximale gegevensoverdracht is veel te laag om zowel digitale JPEG-beelden als digitaal stereogeluid te kunnen bevatten. Er moest dus een nog verdergaande compressietechniek bedacht worden.

Lossy en lossless JPEG

Alvorens deze verdergaande compressietechniek te bespreken moet er nog een opmerking over de JPEG-standaard worden gemaakt. De standaard beschrijft twee compressieschema's:

- Lossy;
- Lossless.

Bij de lossy techniek zal het beeld dat wordt teruggewonnen uit de compressie-

17.2 De MPEG-1 compressie van beeld en geluid

gegevens niet ***identiek zijn aan het originele beeld.*** Er worden dus bij de compressie gegevens verwijderd. Nu valt dat nauwelijks op, omdat het menselijke oog een vrij ongevoelig instrument is. Zo is het oog zeer ongevoelig voor kleine nuances in kleur. Van deze wetenschap wordt bij de lossy techniek gebruik gemaakt om zeer hoge datareducties te verkrijgen. In de tabel van figuur 5/17.2-1 is een overzichtje gegeven van de reductiefactoren die met de lossy techniek haalbaar zijn en de kwaliteit van het herwonnen beeld. Uitgegaan wordt van een "true color" beeld met een kleurendiepte van 24 bit per pixel. Zelfs bij een compressie tot gemiddeld 1,5 bit per pixel blijkt het herwonnen beeld nauwelijks van het origineel te onderscheiden!

BITS PER PIXEL NA COMPRESSIE	COMPRESSIE FACTOR	KWALITEIT NA COMPRESSIE
0,15	160	HERKENBAAR
0,25	96	BRUIKBAAR
0,75	32	GOED
1,50	16	UITSTEKEND
3,00	8	ORIGINEEL

Figuur 5/17.2-1: Uit deze tabel blijkt dat een compressie met een factor 16 een herwonnen beeld oplevert dat nauwelijks van het origineel te onderscheiden is.

Bij de lossless techniek wordt gebruik gemaakt van compressie-algoritmen die tot gevolg hebben dat het originele beeld tot op pixelniveau onaangetast herwonnen kan worden. Het zal duidelijk zijn dat deze techniek veel lagere compressiefactoren oplevert.

MPEG-1

Voor het comprimeren van video en geluid tot een datastroom van maximaal 1,411 Mbit/s werd een nieuwe internationale werkgroep opgericht, die deze keer de naam "Moving Picture Expert Group" kreeg, afgekort tot MPEG. Natuurlijk is het duidelijk dat deze werkgroep voort borduurde op de JPEG-technologie. Het grote verschil tussen JPEG en MPEG is dat JPEG beeldje voor beeldje comprimeert, zonder rekening te houden met de samenstelling van een vorig en/of volgend beeldje. MPEG doet dat wél en bereikt een zeer grote datareductie door alle gegevens die hetzelfde zijn in opeenvolgende beeldjes simpelweg te negeren, of beter uitgedrukt, maar één keer te comprimeren en te verzenden in een zogenoemd referentiebeeld. Nadien worden alleen de verschillen tussen opeenvolgende beeldjes opgezocht en verzonden in verschilbeeldjes. Deze techniek noemt men "motion compensation" en het zal wel duidelijk zijn dat hierdoor een immense datareductie wordt verkregen. Nadat de originele videobeeldjes op deze manier zijn omgezet in weinig (omvangrijke) referentiebeeldjes en veel (minder omvangrijke) verschilbeeldjes worden op sommige beeldjes de JPEG-algoritmen losgelaten. Hierbij heeft men gekozen voor de lossy techniek, omdat het kwaliteitsverlies bij bewegende beelden nog veel minder opvalt dan bij statische beelden. Door een combinatie van de "lossy"-JPEG techniek en de "motion compensation"-MPEG techniek en enige compromissen wat betreft resolutie en kleursamenstelling van het beeld is men in staat video die bestaat uit maximaal 30 beeldjes per seconde te verzenden met een datastroom van maximaal 1,2 Mbit/s. Er blijft dus nog ongeveer 0,2 Mbit/s ruimte over voor het ver-

sturen van het geluid. Wat dat betreft was men snel klaar. Philips (alweer) had immers het indrukwekkende "PASC" ontwikkeld dat gebruikt wordt voor de datareductie van geluid in het DCC-systeem. Dit systeem kon, met enige kleine aanpassingen, zonder meer overgenomen worden in het MPEG-systeem. Ook PASC is een lossy techniek, het herwonnen geluid lijkt (technisch bekeken) in feite helemaal niet op het origineel. Maar ook onze oren zijn een even gebrekkig en onnauwkeurig instrument als onze ogen!

Het ingewikkelde compressieschema van MPEG bleek goed te voldoen voor het doel waarvoor het was ontwikkeld: de consumentenmarkt voorzien van een digitaal videosysteem op CD met een kwaliteit die vergelijkbaar is met deze van de algemeen aanvaarde VHS videorecorder technologie. De techniek leek echter ook veelbelovend voor meer eisende professionele toepassingen. Maar op deze markt was er geen plaats voor de compromissen die waren gesloten. Voor professionele beeldbewerkingstoepassingen moest er dus een nieuwe standaard worden geschapen en vandaar dat de eerste, die voor de consumentenmarkt, MPEG-1 werd gedoopt. De technologie werd beschreven in de ISO/IEC normering 11.172.

Specificaties van MPEG-1

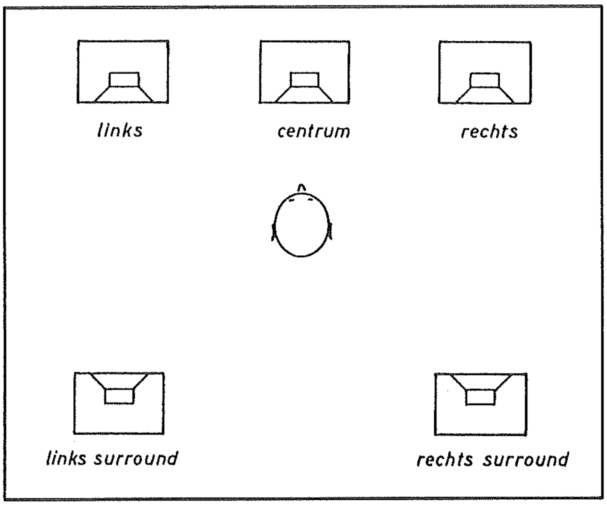
MPEG-1 heeft onderstaande specificaties:

- Ondersteunde systemen:
  PAL, SECAM, NTSC
- Beeldopbouw:
  Niet geïnterlinieerd
- Maximale bitsnelheid video:
  1,15 Mbit/s
- Resolutie luminantie:
  352 bij 288 pixels voor PAL
- Resolutie chrominantie:
  176 bij 144 pixels voor PAL
- Beeldfrequentie:
  30 Hz maximaal, in Europa wordt echter gebruik gemaakt van 25 niet geïnterlinieerde beelden per seconde
- Speelduur:
  Maximaal 72 minuten, een en ander afhankelijk van de noodzakelijke compressiegraad
- Vertraging bij decoderen:
  Maximaal 250 ms tussen het ontvangen van de seriële datastroom van een beeldje en het beschikbaar zijn van het herwonnen beeld
- Audiocodering
  Perceptuele codering van het audiobereik in 32 subbanden met gelijke bandbreedte
- Audiosampling:
  32 kHz, 44,1 kHz, 48 kHz
- Audiosysteem:
  mono, tweekanaals, stereo, joint stereo
- Audio datastroom:
  Kan softwarematig ingesteld worden tussen 32 kbit/s en 448 kbit/s, voor Video-CD begrensd tot 200 kbit/s

MPEG-2

Voor professionele toepassingen werd MPEG-2 ontwikkeld. Deze techniek biedt een veel grotere videoresolutie en mogelijkheden voor het samenvoegen van een groot aantal videokanalen, voor het versleutelen van de gegevens en voor bescherming tegen overdrachtsfouten tot op bitniveau. Bovendien kan MPEG-2 geïnterlinieerde video verwerken, zodat het beter aansluit op de bestaande TV-technologie. Op audioniveau biedt MPEG-2 veel meer bemonsteringsfrequenties en de mogelijkheid tot meerkanaalsgeluid, zoals het door Dolby ontwikkelde en van de bioscoop bekende "surround sound system". Er kunnen vijf audiostromen gecodeerd worden, zie figuur

5/17.2-2, terwijl er dan nog ruimte is voor zeven commentaarkanalen.



Figuur 5/17.2-2: De surround sound mogelijkheden van MPEG-2.

MPEG-2 werd genormeerd in ISO/IEC 13.818 en is in zoverre compatibel met MPEG-1 dat een MPEG-1 datastroom zonder problemen door een MPEG-2 decoder verwerkt kan worden. Een andere belangrijke eigenschap van MPEG-2 is dat er een zogenoemde "low delay"-mode geprogrammeerd kan worden, waarbij de systeemvertraging gereduceerd wordt tot maximaal 100 ms.
Hierdoor wordt het systeem geschikt voor real time bidirectionele communicatie, bijvoorbeeld tussen twee beeldtelefoons. MPEG-2 kent vier formaten, waarvan de eigenschappen zijn samengevat in de tabel van figuur 5/17.2-3.
De audionorm van MPEG-2 zal worden gebruikt voor de op te zetten digitale radio DAB (Digital Audio Broadcasting) waarvan de eerste reguliere uitzendingen in 1995 van start gaan. Verdere toepassingen liggen in digitale satellietverbindingen, zoals "Astra Digital Radio", communicatie via het ISDN-net, abonnee-TV via de satelliet (Hughes Direc-TV), het opvragen van videogegevens volgens de VOD-standaard (Video On Demand) en digitale kabel-TV met 500 kanalen.

Opmerking
Zoals duidelijk blijkt levert zelfs het laagste formaat van MPEG-2 (LDTV) een te grote datastroom op om op een standaard CD-speler verwerkt te kunnen worden. De ontwikkelingen gaan tegenwoordig echter zo snel dat de per definitie traag werkende normeringsinstituten hopeloos achter lopen. Hoe men gaat reageren op het nieuwe verschijnsel "Multi-spin", waarbij CD-spelers de CD op dubbele tot vierdubbele snelheid laten draaien en de gegevens dus twee tot vier maal sneller uitgelezen kunnen worden, is op dit moment nog niet bekend. In principe zou men op een CD-speler die op viervoudige snelheid werkt de datastroom van het SDTV-formaat zonder problemen kunnen verwerken. Dat zou dus prachtige mogelijkheden bieden om ofwel de speelduur te verlengen, zodat een standaard speelfilm op één CD past, of om over te schakelen naar een betere kwaliteit.

FORMAAT	RESOLUTIE	DATASTROOM	KWALITEIT
HDTV	1.920 bij 1.152	30 Mbit/s	ZEER HOOG
EDTV	960 bij 576	10 Mbit/s	BETER DAN PAL
SDTV	720 bij 576	5 Mbit/s	PAL
LDTV	352 bij 288	2 Mbit/s	VHS

Figuur 5/17.2-3: De vier formaten van videobeelden, gecodeerd met MPEG-2.

De algemene structuur van MPEG-1
Bij het comprimeren van video met geluid moet de compressietechniek rekening houden met drie zaken:
- de compressie van de video-informatie;
- de compressie van de audio-informatie;

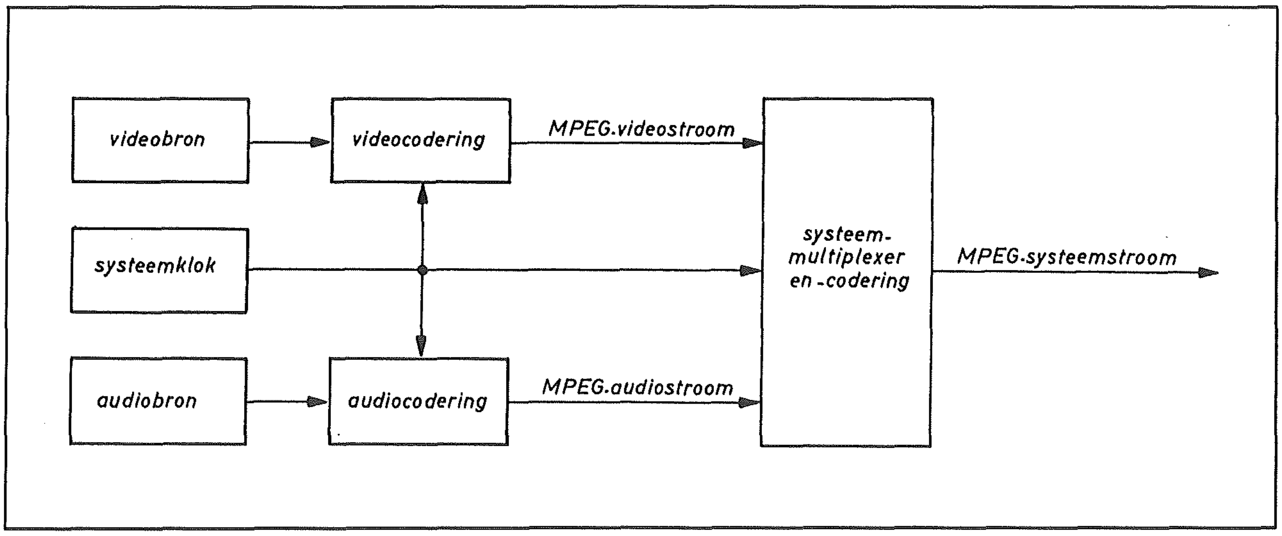
17.2 De MPEG-1 compressie van beeld en geluid

- de synchronisatie tussen video en audio.

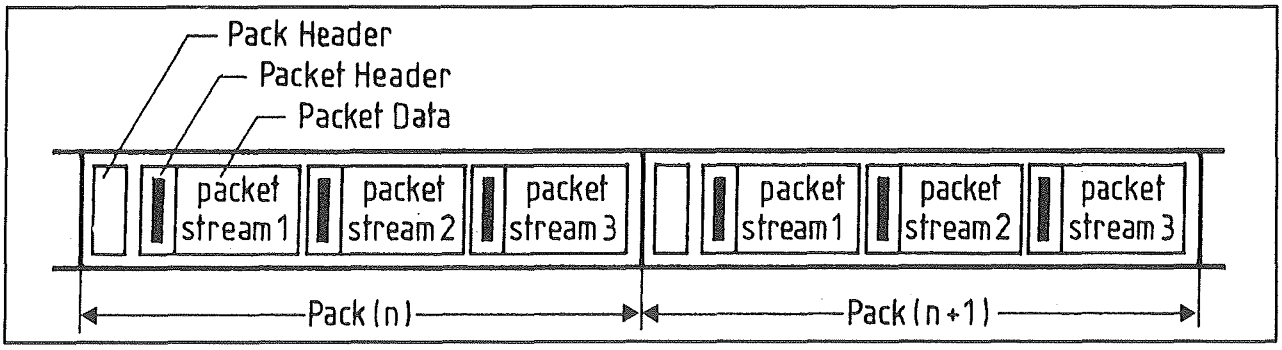
Dit laatste is een belangrijk punt, want omdat de gegevens naar een seriële code worden omgezet is die synchronisatie niet vanzelf sprekend. Toch is het de bedoeling dat de herwonnen gegevens een lipsynchrone film opleveren!

Het algemene schema van een codering volgens MPEG-1 is samengevat in figuur 5/17.2-4. Er zijn twee coders, een voor de video en een voor de audio, een systeemklok en een multiplexer die zorgt dat de gecomprimeerde video- en audiogegegevens op de juiste manier tot een seriële datastroom worden gecombineerd.

De gegevens worden verpakt in blokken die "pack's" worden genoemd. In één pack kunnen video- en audio-gegevens gemengd worden. Ieder pack wordt voorafgegaan door een zogenoemde "header", waarin de voor het decodeersysteem noodzakelijke gegevens over het soort gegevens in het pack en de toegepaste compressie-algoritmen worden verwerkt. Een pack bestaat uit verschillende zogenoemde "packet's", die weer voorafgegaan worden door een subheader en die alleen video- of audiogegevens bevatten. Dat algemene verzendschema is samengevat in figuur 5/17.2-5 en wordt in detail besproken in het volgend hoofdstuk.



Figuur 5/17.2-4: Het algemene schema van een MPEG-1 codering.



Figuur 5/17.2-5: De structuur van de gemultiplexte video- en audiogegevens in de seriële datastroom.

In de packet headers worden "Presentation Time Stamps" opgenomen als tijdreferentie. Deze geven de waarde van de systeemclock weer op het moment dat de gegevens aan de coderingsschakeling worden aangeboden. In de pack header wordt bovendien een "System Clock Reference" opgenomen. Aan de hand van deze gegevens kan de decodeerelektronica de clock herwinnen en kan ervoor gezorgd worden dat de video- en audiogegevens weer volledig worden gesynchroniseerd.

Open standaard

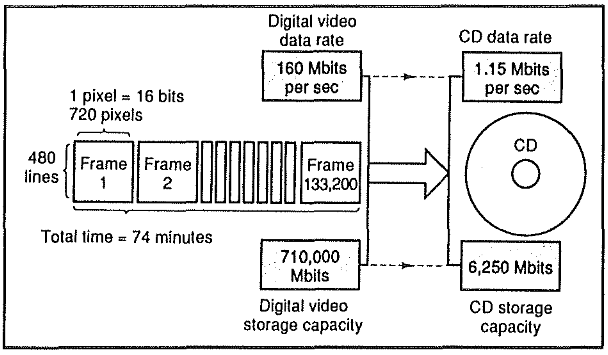
Zoals reeds beschreven heeft MPEG-1 een open structuur. De coderings- en decoderingshardware is in staat verschillende soorten compressieschema's en -algoritmen door te voeren. De gebruikte algoritmen worden gecodeerd en in de headers verzonden. Op deze manier is het mogelijk dat een met MPEG-1 gecomprimeerde film van leverancier X na decodering een betere kwaliteit levert dan dezelfde film, gecodeerd door leverancier Y. De kwaliteit van de herwonnen gegevens wordt immers in hoge mate bepaald door de compressie-algoritmen die worden gebruikt. Die open architectuur maakt het bovendien ook mogelijk dat bepaalde leveranciers over gaan tot het versleutelen van de gegevens. Aan de decoderzijde kan men dan alleen via een gehuurde of gekochte softwaresleutel de gegevens tot een normaal te genieten beeld herleiden. Zelfs de hardware van de codering is in staat tijdens het comprimeren de gebruikte compressie-algoritmen aan te passen. Een voorbeeld. Stel dat een filmfragment wordt aangeboden waar seconden lang nauwelijks beweging in zit, bijvoorbeeld een lang aangehouden close-up van een star gezicht. Het zal duidelijk zijn dat de beeld-tot-beeld compressie door middel van motion compensation dan een erg hoge winst oplevert. De hardware kan dan besluiten deze gegevens door een minder ver gaand lossy JPEG-algoritme te laten comprimeren. De kwaliteit van de herwonnen beelden is dan erg hoog. Als even later het beeld echter vol actie komt zal de motion compensation veel minder gegevensreductie opleveren. Zou nu hetzelfde percentage lossy compressie worden toegepast, dan zou de datastroom de magische grens van 1,411 Mbit/s overschrijden. De hardware beslist dan zelfstandig dat het noodzakelijk is een grotere compressiefactor in de lossy techniek van JPEG toe te passen. De kwaliteit van de herwonnen beelden gaat dan dalen, maar gelukkig is het oog daar op dat moment allerminst gevoelig voor. Het oog neemt namelijk voornamelijk de actie op het scherm waar en let niet op details.

De MPEG-1 codering van video

Inleiding

De zware taak, waar de bedenkers van het systeem voor gesteld stonden, wordt dramatisch toegelicht aan de hand van figuur 5/17.2-6. De analoge beelden van bijvoorbeeld een film worden een na een gescand met een resolutie van 480 lijnen die ieder 720 pixels bevatten. Dit is in ieder geval de Amerikaanse norm. Per pixel worden geen 3 x 8, maar slechts 16 bit gebruikt (zie later). Per beeld levert dat ongeveer 5,5 Mbit aan gegevens op. Als men 25 beelden per seconde wil versturen heeft men daarvoor ongeveer 138 Mbit nodig! Natuurlijk moet men nog allerlei

coderings- en compressiegegevens meesturen, zodat één seconde video een datastroom van ongeveer 160 Mbit/s oplevert. De CD kan echter (afgerond naar een veilige ondermarge en rekening houdende met ruimte voor de audio) slechts ongeveer 1,15 Mbit/s verwerken. De gegevens moeten dus met een factor 140 gecomprimeerd worden!



Figuur 5/17.2-6: Uit dit overzichtje blijkt dat de videogegevens met een factor 140 gecomprimeerd moeten worden.

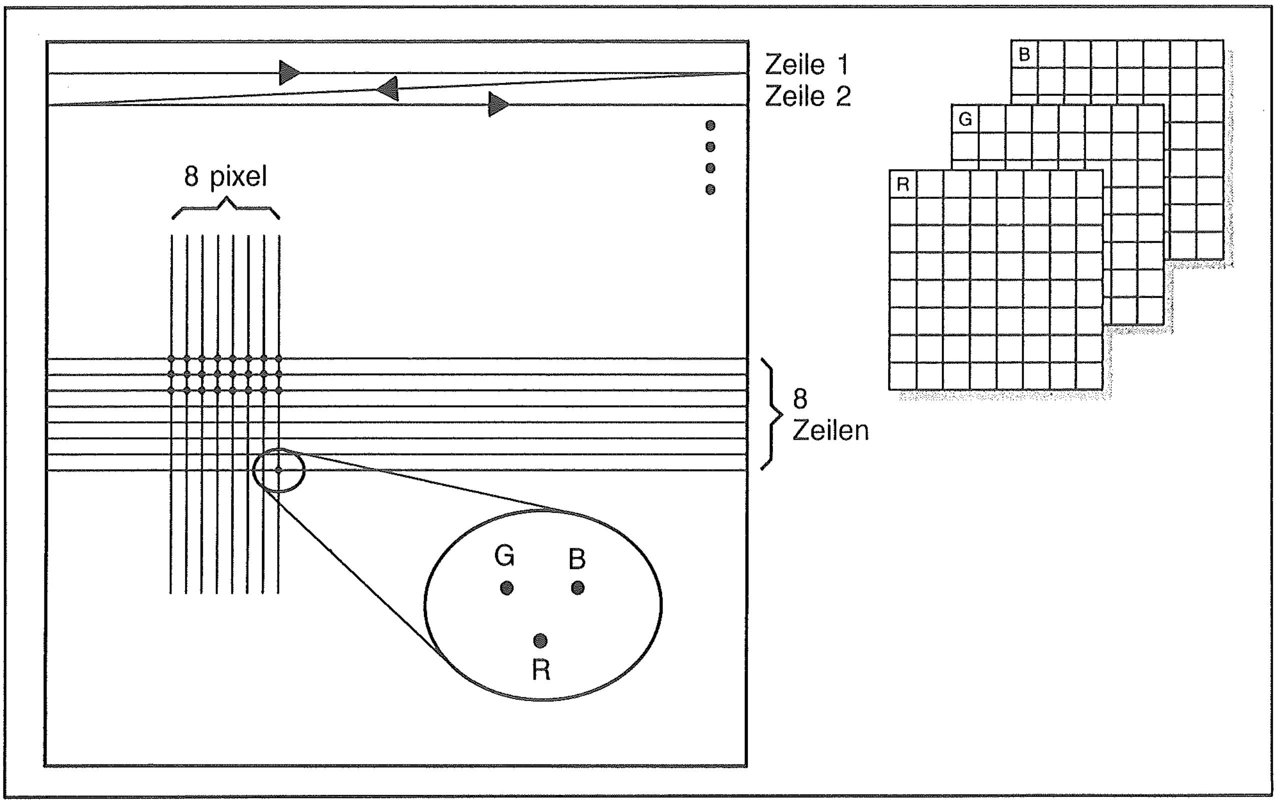
Het zal duidelijk zijn dat een dergelijke zware compressie niet in één stap te realiseren is. De codering van de videogegevens is erg ingewikkeld en doorloopt een heleboel stapjes. In de meeste stappen wordt iets bijgedragen aan de totale compressie, zodat uiteindelijk toch de noodzakelijke totale compressiefactor verkregen wordt.

Overzicht

In het kort komt de MPEG-verwerking van een analoog filmsignaal naar de uiteindelijke gecomprimeerde datastroom neer op de volgende procedure:

- Scannen van de filmbeelden in 3 x 8 bit volgens het RGB-kleurenschema en met een hoge resolutie van 720 bij 480.
- Intelligent reduceren van de horizontale en vertikale resolutie met een factor twee (compressieslag 1).
- Omzetten van de RGB-gegevens naar een YUV-schema en intelligente reductie van de chrominantie-gegevens (compressieslag 2).
- Indelen van het beeld in zogenoemde macroblokken van 8 bij 8 pixels.
- Opslaan van een groot aantal macroblokken van opeenvolgende beeldjes in een geheugen en toepassen van motion compensation door middel van forward prediction en bidirectional prediction. Verdelen van de beeldenstroom in zogenoemde I-, P- en B-frames, berekenen van zogenoemde "motion vectors" en berekenen van de zogenoemde "prediction errors" (compressieslag 3).
- Toepassen op de macroblokken van de I-beelden van de zogenoemde DCT, oftewel de "discrete cosinus transformatie".
- De gegevens die hieruit voortkomen onderwerpen aan een zogenoemde "psychovisuele quantiseringsmatrix", waardoor de lossy compressie van JPEG ontstaat (compressieslag 4).
- De gequantiseerde gegevens van een macroblok zigzaggend omzetten in een seriële datastroom.
- De JPEG-gecomprimeerde gegevens van de I-frames weer combineren met de gegevens van de P- en B-frames en deze in een bepaalde volgorde samenvoegen tot één datastroom.
- De seriële gegevens onderwerpen aan een zogenoemde "run length codering" (compressieslag 5).
- De omgezette gegevens comprimeren door middel van een zogenoemde "entropie-codering" op basis van het Huffmann-algoritme (compressieslag 6).

17.2 De MPEG-1 compressie van beeld en geluid



Figuur 5/17.2-7: Het scannen van de filmbeeldjes van een analoge film.

Een heleboel nieuwe begrippen, die in de volgende paragraafjes een na een hun geheimen prijs zullen geven!

Het scannen

Bij de toepassing waar het hier om gaat, Video-CD, zal het bronmateriaal in de meeste gevallen bestaan uit een "ouderwetse" analoge film. Een film samengesteld uit tienduizenden dia's op een rol die via een sluiter en transportmechanisme met een snelheid van 24 beeldjes per seconde op het witte doek worden geprojecteerd. Er bestaan reeds jaren filmscanners, die beeldje na beeldje omzetten in een digitaal beeld. Deze scanners werken met een resolutie van 720 bij 576 bij PAL-televisie en met 720 bij 480 bij het Amerikaanse NTSC-systeem. Dat wil zeggen dat ieder filmbeeld, zie figuur 5/17.2-7, in 576 lijnen wordt afgetast en dat er in iedere lijn 720 beeldmonstertjes worden genomen. Uit de aard van hun werking leveren scanners een zogenoemd "RGB-schema" af. Ieder monster wordt ontleed in de drie basiskleuren rood, groen en blauw. Afhankelijk van de kleurverzadiging en de intensiteit wordt aan iedere basiskleur een code tussen "000" en "255" toegekend. Dat levert in totaal een kleurenspectrum van meer dan 16 miljoen kleuren op, het zogenoemde "true color".

Op deze manier wordt het beeld omgezet in drie zogenoemde "planes", een voor rood, een voor groen en een voor blauw.

Het reduceren van de resolutie

Het zou erg fraai zijn als de beelden verder verwerkt konden worden in die hoge resolutie van 720 bij 576. Dat kan helaas niet, de resulterende datastroom zou te

groot zijn voor een CD. Vandaar dat de resolutie wordt gereduceerd tot 352 bij 288. Dat gebeurt niet door eenvoudigweg ieder tweede pixel te laten wegvallen. In horizontale richting wordt een intelligent reductieschema toegepast, waarbij de elektronica gemiddelde kleurwaarden berekent van de pixels die behouden blijven en de pixels die worden verwijderd. In vertikale richting kan een identiek proces worden toegepast, maar het is ook toelaatbaar dat simpelweg lijnen worden geschrapt.

De RGB naar YUV vertaling

Een kleurenbeeld kan niet alleen volledig worden beschreven door de R-, G- en B-waarden. Het is ook mogelijk een beeld te definiëren door de luminantie Y en de chrominantie C. De eerste grootheid geeft informatie over de helderheid van het beeld, de tweede bepaalt de kleursamenstelling. Dit schema wordt bijvoorbeeld bij kleuren-TV gebruikt. Op een zwart/wit televisie wordt alleen de Y-waarde gebruikt. Om de kleur te omschrijven moet men wel twee grootheden voor de chrominantie definiëren. Deze worden C_U en C_V genoemd. Het grote voordeel van een definiëring volgens het YUV-schema is dat luminantie en chrominantie elkaar niet beïnvloeden en volledig afzonderlijk verwerkt kunnen worden. Het menselijke oog blijkt heel gevoelig voor variaties in de luminantie, maar veel minder gevoelig voor variaties in de chrominantie. Vandaar dat men kan volstaan met een kleinere resolutie bij het definiëren van de chrominantie.

De verbanden tussen de R-, G- en B-waarden en de Y-, C_U- en C_V-waarden volgt uit de matrix van figuur 5/17.2-8. Voor wie niet zo vertrouwd is met wiskundige formuleringen onder de vorm van een matrix worden onderstaand de transformatieformules voluit geschreven:

$Y = 0,299.R + 0,587.G + 0,114.B$

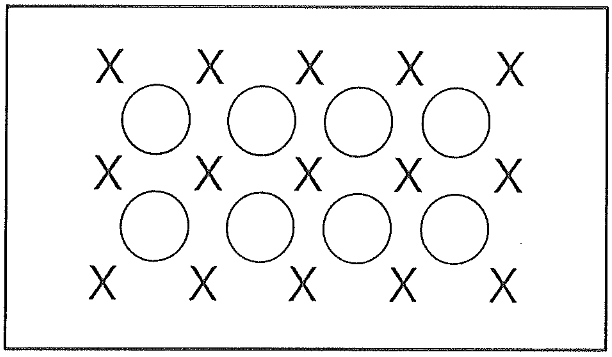
$C_U = -0,169.R - 0,3316.G + 0,5.B$

$C_V = 0,5.R - 0,4186.G - 0,0813.B$

Voor het omrekenen moet de processor van de coder dus slechts enige eenvoudige rekenkundige bewerkingen op de R-, G- en B-waarden uitvoeren, een fluitje van een cent. Er ontstaan nu drie nieuwe "planes", waarbij de "plane" van Y nog steeds een resolutie van 352 bij 288 heeft. Omdat de informatie in de chrominantie niet zo belangrijk is, wordt de resolutie daarvan dramatisch verminderd. Er wordt een gemiddelde waarde berekent per vier pixels, volgens het schema van figuur 5/17.2-9.

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.3316 & 0.500 \\ 0.500 & -0.4186 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Figuur 5/17.2-8: De matrix die het verband geeft tussen de RGB-definitie van een kleurenbeeld en de YUV-definitie.



Figuur 5/17.2-9: Het reduceren van de chrominantiegegevens.

De kruisjes stellen de pixels van het gedigitaliseerde beeld voor. Van ieder pixel wordt eerst de luminantiewaarde bere-

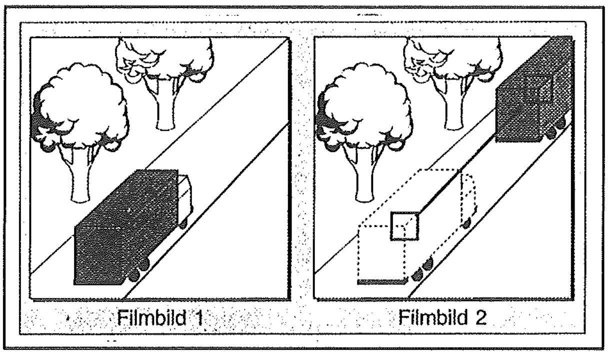
kend. Nadien worden de twee chrominatiewaarden van vier naast elkaar gelegen pixels berekend en wordt er de gemiddelde waarde van bepaald. Deze gemiddelde waarde wordt voorgesteld door de cirkeltjes. De planes van de twee C-waarden zijn door de reductie nu nog slechts 176 bij 144 pixels groot, alweer een aanzienlijke reductie waar de kwaliteit van het beeld nauwelijks door beïnvloed wordt!

Het maken van macroblokken

Voor de daaropvolgende ingewikkelde wiskundige bewerkingen zijn de volledige planes veel te groot. Deze worden ingedeeld in zogenoemde "macroblokken", die slechts 8 bij 8 pixels groot zijn. De volledige definitie van een stukje beeld van 16 bij 16 pixels, op een standaard beeldbuis goed voor ongeveer 120 vierkante millimeter beeldoppervlak, bestaat dus uit vier macroblokken voor de luminantie en twee macroblokken voor de chrominantie. Vervolgens worden een heleboel van dergelijke macroblokken in een geheugen opgeslagen. Niet alleen alle macroblokken van het beeld dat verwerkt wordt, maar ook deze van enige voorgaande en volgende beelden.

De motion compensation

Deze compressieslag vormt het hart van de MPEG-technologie. Het filmbeeld bestaat uit 25 beeldjes per seconde. Nu zal het duidelijk zijn dat er in opeenvolgende filmbeeldjes veel identieke informatie zit. Stel als voorbeeld het filmfragment dat in figuur 5/17.2-10 is voorgesteld. Een vrachtwagen rijdt van links onder naar rechts boven door het beeld. De achtergrond, bomen langs de weg, blijft hetzelfde. Het is dan overbodig om deze video-informatie in ieder beeld te gaan coderen door middel van de YUV-waarden. Dat kan in één beeld, in de volgende beelden wordt op de een of andere manier duidelijk gemaakt dat de informatie in de betreffende macroblokken ongewijzigd is gebleven. Ook de macroblokken die de vrachtwagen voorstellen veranderen inhoudelijk niet. Het enige verschil is dat de macroblokken in het beeld een andere positie innemen. Het rechter beeld kan dus voor een groot deel samengesteld worden uit de macroblokken van het linker beeld. Natuurlijk zijn er ook elementen in het beeld die nieuw zijn. Daar waar in het linker beeld de vrachtwagen stond in nu een stuk weg plus omgeving te zien. Deze macroblokken bevatten dus beeldspecifieke informatie en moeten wel gecodeerd worden.



Figuur 5/17.2-10: In deze twee beeldjes zit veel identieke informatie, die slechts één keer gecodeerd moet worden.

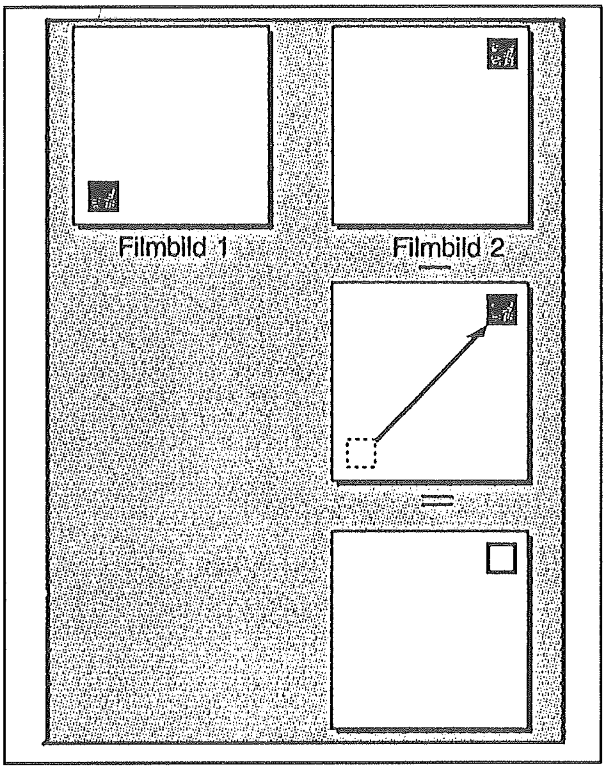
Het principe van de reductie door middel van motion compensation berust op twee grootheden: een verplaatsingsvector, "motion vector" genoemd en een foutcorrectie, "prediction error" genoemd.
Deze begrippen worden toegelicht aan de hand van figuur 5/17.2-11. Stel dezelfde beeldvolgorde als in figuur 5/17.2-10. Een macroblok dat een deel van de achterkant van de vrachtwagen voorstelt is verplaatst van de linker onderhoek van het beeld

naar de rechter bovenhoek van het volgende beeld. Die verplaatsing kan aangegeven worden door een verplaatsingsvector (motion vector), het pijltje in de middelste figuur. Die vector kan gedefinieerd worden door slechts twee getallen: het eerste getal geeft de verplaatsing in horizontale richting weer, het tweede getal de verplaatsing in vertikale richting. Die twee getallen nemen uiteraard heel wat minder ruimte in beslag dan de volledige YUV-codering van het 8 bij 8 pixels grote macroblok! Nu kan het gebeuren dat er toch iets aan de inhoud van het verplaatste macroblok wijzigt. Stel bijvoorbeeld dat het macroblok niet een stuk vrachtwagen, maar een stuk van de jas van een bewegende man voorstelt.
Door schaduwwerking en lichtinval zou het kunnen gebeuren dat de algemene kleurinvulling van het blok wel identiek is, maar dat er iets andere tinten aanwezig zijn. De hardware is in staat dat te detecteren en de prediction error te berekenen. Dat zijn foutgegevens, die de YUV-gegevens van het blok in het linker beeld corrigeren naar de YUV-gegevens van het identieke blok in het rechter beeld. Maar deze foutcorrectie kan veel effectiever gecodeerd worden dan een volledige pixel na pixel beschrijving van het rechter macroblok! Alweer een niet geringe datareductie! In het getekende voorbeeld verandert er niets in de inhoud van het macroblok. De prediction error (onderste figuur) bevat dus geen gegevens, hetgeen wordt voorgesteld door een codering met nullen. In een volgende compressieslag kunnen deze 64 nullen heel eenvoudig worden gecomprimeerd.

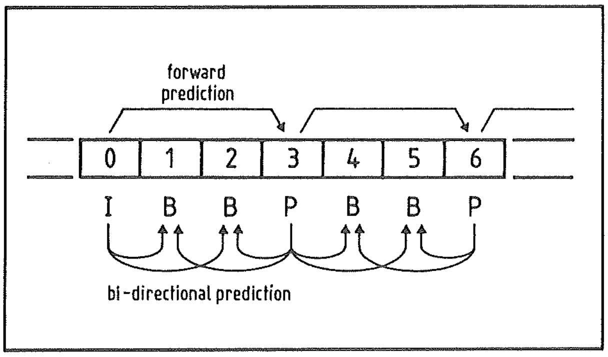
I-, P- en B-beelden

Het zal duidelijk zijn dat het principe van motion vectors maar heel even werkt. Na enige beeldjes is de vrachtwagen bijvoorbeeld volledig uit het beeld verdwenen en kan het systeem zich suf zoeken naar identieke macroblokken. Ook als een plotselinge beeldwisseling optreedt gaat het systeem de mist in. Dit probleem wordt opgevangen door regelmatig referentiebeelden op te nemen. Dat zijn beelden waarop geen compressie door middel van motion vectors en prediction errors wordt toegepast, maar die volledig worden beschreven door de YUV-waarden van al hun pixels. Deze beelden worden "I-frames" genoemd, de afkorting van "Intra frame". Als de scène rustig is worden er gemiddeld twee I-frames per seconde aangemaakt. Bij scènes met veel actie kan die frequentie echter behoorlijk oplopen. Uit één zo'n I-frame kunnen diverse "P-frames" worden afgeleid. Dit proces noemt men "forward prediction", het voorspellen van de inhoud van een beeld uit de gegevens van een vorig beeld. De "P" van P-frame is uiteraard de afkorting van "predicted", zo genoemd omdat de samenstelling van deze frames voor een groot deel uit de inhoud van het I-frame kan worden afgeleid. Het berekenen van de motion vectors is echter voor zelfs de snelste processoren een immense klus. Zou men dat beeldje voor beeldje doen, dan zou het coderen van een film met JPEG-1 veel te veel tijd kosten. Vandaar worden er slechts om de drie beelden P-frames berekend.
De tussenliggende beeldjes worden op een andere manier gecodeerd. Deze worden "B-frames" genoemd, de afkorting van "bidirectional". Deze worden zo genoemd omdat zij worden gecodeerd door interpolatie van de gegevens uit I- en P-frames die links en rechts van het B-frame liggen. Dit proces wordt uiteraard "bidirectional prediction" genoemd.

17.2 De MPEG-1 compressie van beeld en geluid

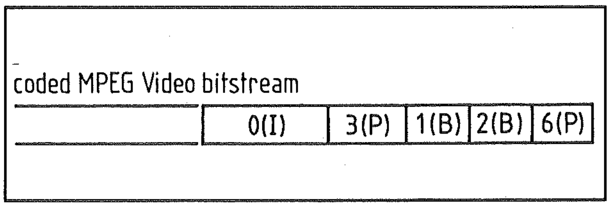


Figuur 5/17.2-11: Het coderen van macroblokken door midden van een motion vector en een prediction error.



Figuur 5/17.2-12: Het omzetten van de normale YUV-beeldjes in opeenvolgende I-, B- en P-frames.

Om nog eens terug te komen op het voorbeeld met de vrachtwagen, het zal duidelijk zijn dat het in de praktijk nooit voorkomt dat een vrachtwagen in één beeldovergang volledig door het beeld rijdt. Tussen de twee in figuur 5/17.2-10 getekende beeldjes liggen er een aantal, waarin de vrachtwagen steeds meer van links naar rechts en van onder naar boven opschuift. Als men nu het linker beeldje als I-frame codeert en het rechter beeldje als P-frame, dan kunnen de tussenliggende beelden gecodeerd worden door de horizontale en vertikale waarden van de verplaatsingsvector van een macroblok te delen door het aantal tussenliggende beeldjes en nadien door middel van stapsgewijze vermeerdering de verplaatsingsvectoren van dat macroblok te coderen in de tussenliggende B-frames. Dat scheelt een heleboel rekentijd voor de processor.
Het algemene schema van de omzetting van de "normale" YUV-gecodeerde beeldjes in een IBP-codering is getekend in figuur 5/17.2-12.
De pijltjes in deze tekening geven de richting aan waarin de processen van forward en bidirectional prediction verlopen. Het zal duidelijk zijn dat de frames echter niet in de getekende volgorde in de seriële datastroom kunnen worden opgenomen. De decoder in de Video-CD speler moet immers eerst over de gegevens van I- en P-frames beschikken, alvorens de inhoud van de B-frames berekend kan worden. Vandaar dat de gegevens door elkaar gehusseld worden en wel volgens het in figuur 5/17.2-13 getekend schema. Eerst wordt het I-frame in de datastroom gezet, dan het eerstvolgende P-frame en vervolgens de tussenliggende B-frames.



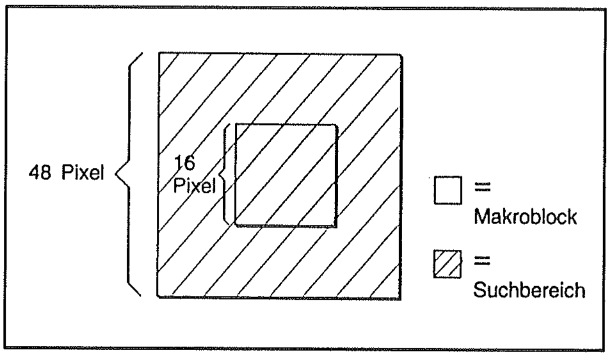
Figuur 5/17.2-13: De volgorde van de I-, P- en B-frames in de seriële datastroom.

Het berekenen van de motion vectors

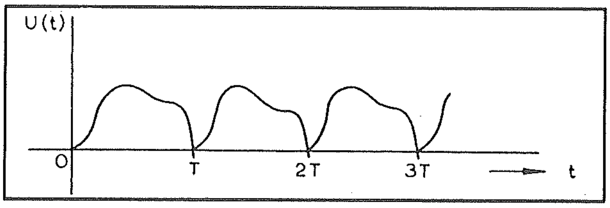
Het berekenen van alle motion vectors is een ongelooflijk ingewikkelde opdracht voor de processor. In de meeste gevallen wordt volgens de volgende systematiek gewerkt. Vier naast elkaar liggende macroblokken uit een I-frame worden gekombineerd tot een 16 bij 16 pixel blok. Van dit blok worden een soort van gemiddelde waarden van de luminantie en de chrominantie berekend. Vervolgens wordt dit blok, zie figuur 5/17.2-14, geplaatst in een groter omliggend gedeelte van het P-frame dat 48 bij 48 pixels groot is. De processor gaat nu in dat groter blok steeds opnieuw blokjes van 16 bij 16 onderzoeken op de soort gemiddelde waarden van luminantie en chrominantie. Als die waarden in het basisblok en het kandidaatblok ongeveer overeenstemmen, gaat het systeem er van uit dat het een identiek blok gevonden heeft in het P-frame en kan het de motion vector berekenen. Nu lijkt het niet te geloven dat een dergelijk systeem in de praktijk werkt. Toch doet het dat bijna feilloos, wat mede te danken is een aan forse hap hogere wiskunde die wordt toegepast. Er werd niet voor niets geschreven dat de processor "een soort" gemiddelde waarden berekent. In de praktijk worden de wiskundige wetten van de stochastiek, de waarschijnlijkheidsleer, uit de kast gehaald om de waarschijnlijke gelijkheid van blokken te beoordelen. Iemand die iets van dit soort wiskunde afweet zal het begrip "gemiddelde kwadratische differentieelwaarde" misschien iets zeggen!

De discrete cosinus transformatie

In de vorige bewerking werd reeds een heleboel hogere wiskunde toegepast. Bij de DCT, oftewel discrete cosinus transformatie, worden de allerhoogste wiskundige registers open getrokken! DCT is een proces, waarbij een tweedimensioneel continu amplitudeverloop kan omgezet worden in een ruimtelijk discreet frequentiespectrum. Dat zal de gemiddelde niet-wiskundig geschoolde lezer helemaal niets zeggen. Omdat de DCT het absolute hart van MPEG en JPEG is, is het toch belangrijk dat dit proces begrijpelijk wordt gemaakt. Dat kan door een ommetje te maken naar een begrip dat de meeste elektronici wél iets zal zeggen: de Fourier-analyse.



Figuur 5/17.2-14: Het doorzoeken van blokken van 48 bij 48 pixel uit een volgend beeld op dezelfde "gemiddelde waarden" van luminantie en chrominantie als in het 16 bij 16 pixel blok uit het voorgaande beeld.

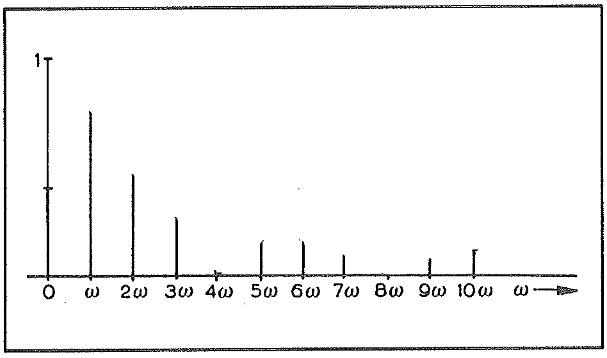


Figuur 5/17.2-15: Een periodiek analoog signaal dat door de Fourier-analyse kan omgezet worden in een frequentiespectrum.

In figuur 5/17.2-15 is een periodiek analoog signaal voorgesteld. Dit signaal wordt in deze figuur beschreven door het conti-

17.2 De MPEG-1 compressie van beeld en geluid

nue verloop van zijn momentele waarde (het amplitudeverloop) te tekenen in een assenstelsel, waarin de amplitude U(t) wordt uitgezet in functie van de tijd t. Dergelijke signalen zijn, hoe gek het ook klinkt, samengesteld uit niets meer dan een bepaalde gelijkspanning en verder een heleboel zuivere sinusoidale wisselspanningen met frequenties die gelijk zijn aan veelvouden van de eigen frequentie van het signaal. Iedere frequentie heeft een eigen grootte en de verhouding tussen deze groottes definieert volledig de vorm van het signaal. Deze zuiver sinusvormige spanningen noemt men "de harmonischen" van het signaal. Het signaal van figuur 5/17.2-15 kan dus ook getekend worden onder de vorm van het frequentiespectrum van figuur 5/17.2-16.



Figuur 5/17.2-16: Het frequentiespectrum van het signaal van figuur 5/17.2-15.

Dat is een lijndiagram, waarbij iedere lijn staat voor een harmonische. Deze worden gedefinieerd door hun frequentie op de horizontale frequentie-as (ω) en door hun grootte op de vertikale amplitude-as (1). De ene gelijkspanning noemt men de DC-component van het signaal, alle sinusspanningen de AC-componenten. Als men een mengschakelingetje zou bouwen en aan de ingangen van deze menger een gelijkspanning en alle harmonischen zou aanbieden, dan kan men de mengpotentiometers zo instellen dat het signaal van figuur 5/17.2-15 uit de uitgang van de menger komt. Het vreemde van de Fourier-analyse is dat de grootte van de harmonischen afneemt naarmate hun frequentie stijgt. Dat heeft als gevolg dat die hoge harmonischen in feite niet erg veel bijdragen aan de vorm van het signaal. Zou men bijvoorbeeld de potentiometers van de signalen 9.ω en 10.ω op nul draaien, zodat deze signalen niet in het mengproces betrokken worden en zou men met een oscilloscoop het signaal op de uitgang van de menger observeren, dan zou men vast stellen dat het signaal er nauwelijks anders zou uitzien. Nu terug naar de discrete cosinus transformatie! Maar even een recapitulatie. De beeldenstroom die werd aangevoerd als YUV-gecodeerde frames is in het vorig proces omgezet in een stroom van I-, B- en P-frames. Alleen de I-frames bevatten echte YUV-informatie, de B- en P-frames bevatten gecodeerde motion vectors en prediction error codes. De I-frames worden nu onderworpen aan een discrete cosinus transformatie. Dat geldt zowel voor de luminantie- als voor de chrominantiegegevens. Bekijk even (figuur 5/17.2-17) de samenstelling van een macroblok uit een I-frame, dat de informatie over de luminantie van het blok bevat. Die informatie kan voorgesteld worden door een matrix met 64 cellen waarbij iedere cel de luminantiewaarde van het betreffende pixel bevat. Dat is een getal tussen "000" en "255". In feite kan deze informatie ook ruimtelijk voorgesteld worden. Dit is getekend in het ruimtelijke diagram van figuur 5/17.2-18. Het grondvlak stelt de positie van de pixels uit het macroblok voor, de hoogte-as de waarde van de luminantie. De toppen van de staven vormen

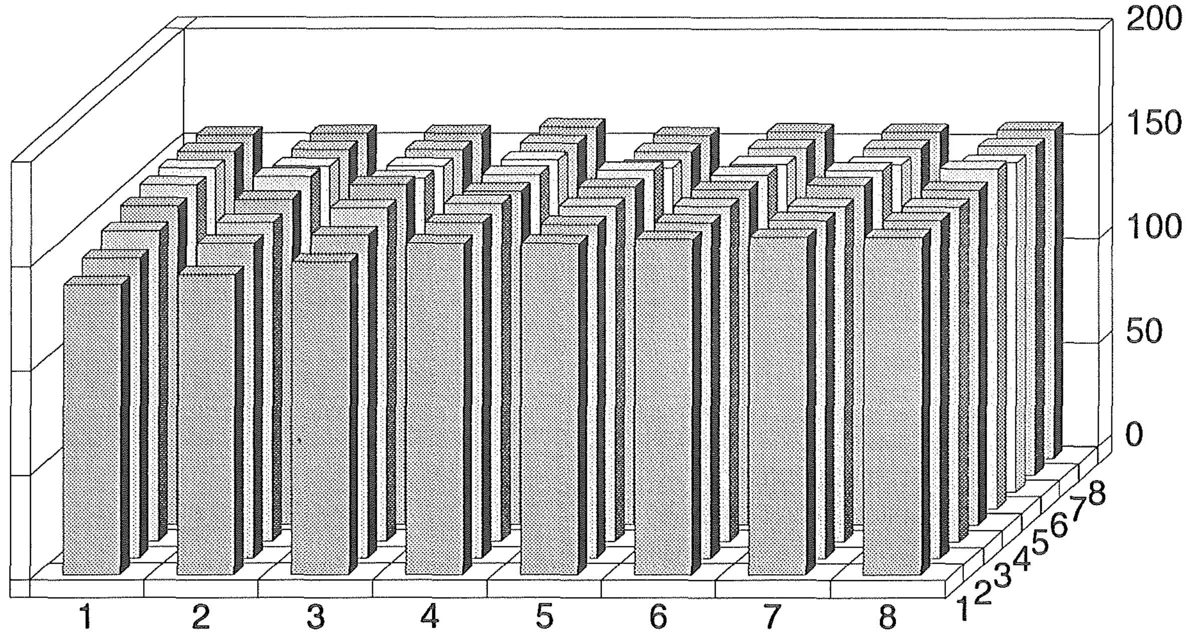
een glooiend patroon, op te vatten als een heuvelachtig landschap. De heuvels stellen het verloop van de luminantie voor in het macroblok. Dit is dus in feite niets anders dan het amplitudeverloop van het signaal van figuur 5/17.2-15, maar nu in een tweedimensioneel vlak in plaats van langs één as. Dit amplitudeverloop verloopt vrij vlak en dat is geen wonder want het gaat over niets meer dan een heel klein deeltje van een beeld, namelijk een gebied van 8 bij 8 pixels. De vraag kan gesteld worden of het mogelijk is op deze ruimtelijke amplitudeverdeling ook een soort Fourier-analyse toe te passen, dus het amplitudeverloop om te zetten in een ruimtelijk frequentiespectrum. Dat kan inderdaad, maar niet met de Fourier-analyse maar met een zeer ingewikkelde wiskundige theorie, die discrete cosinus transformatie wordt genoemd. Voor de liefhebbers van formules worden in figuur 5/17.2-19 de vier transformatieformules gegeven. De eerste formule voert de transformatie in de ene richting uit, de derde in de andere richting. De twee overige formules kunnen gebruikt worden om uit de getransformeerde gegevens weer de originele luminantiewaarden te berekenen. Het gevolg van de DCT is dat er een nieuwe matrix ontstaat met nog steeds 64 getalletjes. Voor de matrix van figuur 5/17.2-17 levert dit het resultaat van figuur 5/17.2-20 op (reken maar na!). Die getallen, die dus ieder de grootte van een harmonische geven, worden volgens een zigzag-schema in de matrix geschreven (figuur 5/17.2-21).
In de linker bovenhoek staat het getal van de DC-component (1.260), rechts daarnaast de amplitude van de eerste harmonische (-1), links onder de amplitude van de tweede harmonisch (-23). Dat gaat al zigzaggend zo verder tot in de rechter benedenhoek de amplitude van de hoogste harmonische (0) wordt genoteerd.

139	144	149	153	155	155	155	155
144	151	153	156	159	156	156	156
150	155	160	163	158	156	156	156
159	161	162	160	160	159	159	159
159	160	161	162	162	155	155	155
161	161	161	161	160	157	157	157
162	162	161	163	162	157	157	157
162	162	161	161	163	158	158	158

Figuur 5/17.2-17: De getallenmatrix, die de luminantie-informatie van een macroblok uit een I-frame bevat.

De voordelen van DCT
Het toepassen van een discrete cosinus transformatie heeft een aantal voordelen. Maar allereerst moet een misverstand uit de wereld worden geholpen. De matrix van figuur 5/17.2-17 kan geïnterpreteerd worden als het reële verloop van de luminantie over een beeldbereik van 8 bij 8 pixels. De getallen in de DCT-matrix van figuur 5/17.2-20 kunnen niet als dusdanig geïnterpreteerd worden. Het zijn 64 getallen, die op een volledig wiskundige manier het luminantieverloop beschrijven door aan te geven hoe het ruimtelijk frequentiespectrum van het macroblok er uit ziet. Let nog steeds op de overeenkomst met Fourier-analyse van een elektrisch signaal!

17.2 De MPEG-1 compressie van beeld en geluid



Figuur 5/17.2-18: De luminantieverdeling in een macroblok, voorgesteld onder de vorm van een amplitudeverloop.

De DCT-omzetting heeft als eerste groot voordeel dat het systeem minder gevoelig wordt voor fouten in de decodering. Stel dat de Video-CD speler een heel foutieve waarde uit de spiraal uitleest. Zou men zonder meer luminantie- en chrominantiewaarden in de spiraal coderen, dan zou ieder foutief uitgelezen bit een felle lichtflits op het scherm tot gevolg kunnen hebben. Vergeet niet dat het formaat van Video-CD (zie volgend hoofdstuk) niet erg veel ruimte overlaat voor een grondige foutcorrectie! Leest de speler echter een foutieve waarde van de DCT-matrix uit, dan zal dit alleen tot gevolg hebben dat het luminantieverloop over een heel gebied van 8 bij 8 pixels iets anders wordt. Dat valt veel minder op!

Het tweede grote voordeel van de DCT is dat, onderzoek de DCT-matrix van figuur 5/17.2-20, er een heleboel hogere harmonischen zijn die een erg kleine amplitude hebben. Alweer vergelijkbaar met de Fourier-analyse! In feite zit de belangrijkste beeldinformatie nu alleen in de linker bovenhoek van de matrix. Een heel groot verschil met de grote getallen over de gehele matrix van figuur 5/17.2-17! Als al die kleine amplitudes van de hogere harmonischen op nul worden gesteld, zal dat nauwelijks gevolgen hebben voor het verloop van de luminantie in het macroblok (denk aan het verhaal van de menger!). De DCT bereidt het systeem dus voor op een grote compressieslag, die in de volgende fase tot stand komt.

17.2 De MPEG-1 compressie van beeld en geluid

Formel 1: 2D FDCT

$$C(u,v) = \frac{1}{4} C_u C_v \sum_{x=0}^{7} \sum_{y=0}^{7} f(x,y) \cos \frac{(2y+1)u\pi}{16} \cos \frac{(2x+1)u\pi}{16}$$

Formel 2: 2D IDCT

$$f(x,y) = \frac{1}{4} \sum_{u=0}^{7} \sum_{v=0}^{7} C_u C_v C(u,v) \cos \frac{(2y+1)u\pi}{16} \cos \frac{(2x+1)v\pi}{16}$$

$C_u, C_v = \frac{1}{\sqrt{2}}$ für u,v = 0 $\quad C_u, C_v = 1$ für u,v = 1...7

Formeln der 1D DCT

Formel 3: 1D FDCT

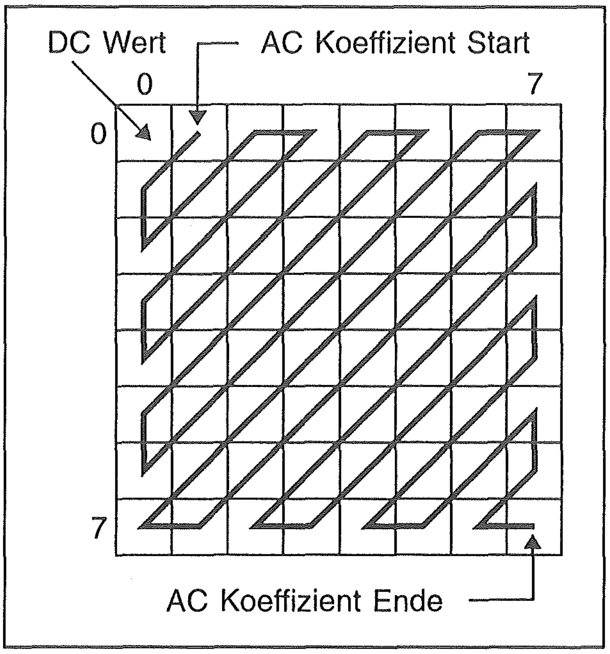
$$C(u) = \frac{1}{2} C_u \sum_{x=0}^{7} f(x) \cos \frac{(2x+1)u\pi}{16}$$

Formel 4: 1D IDCT

$$f(x) = \frac{1}{2} \sum_{u=0}^{7} C_u C(u) \cos \frac{(2x+1)u\pi}{16}$$

$C_u, C_v = \frac{1}{\sqrt{2}}$ für u = 0 $\quad C_u = 1$ für u = 1...7

Figuur 5/17.2-19: Wiskundige magie: de formules voor het omzetten van het "amplitudeverloop" van de luminantie in harmonische frequenties met specifieke amplitudes.



Figuur 5/17.2-21: De zigzaggende manier waarop de frequentie-amplitudes in de DCT-matrix worden geschreven.

1260	-1	-12	-5	2	-2	-3	1
-23	-17	-6	-3	-3	0	0	-1
-11	-9	-2	2	0	-1	-1	0
-7	-2	0	1	1	0	0	0
-1	-1	1	2	0	-1	1	1
2	0	2	0	-1	1	1	-1
-1	0	0	-1	0	2	1	-1
-3	2	-4	-2	2	1	-1	0

Figuur 5/17.2-20: De getallenmatrix die ontstaat na de discrete cosinus transformatie.

De psychovisuele quantisering van de DCT-matrix

Quantiseren is alweer een wiskundig begrip, dat simpelweg vertaald er op neer komt dat grote getallen worden omgezet in kleinere getallen door de grote getallen te delen door een bepaalde factor. Kleine getallen worden door deze deling tot nul gereduceerd. Als men dus de gegevens uit een DCT-matrix quantiseert, dan zullen er een heleboel nullen in de matrix ontstaan. Dit is de basis van de lossy compressie van JPEG. DCT-waarden die door de quantisering worden gereduceerd tot nul, kunnen namelijk nooit ofte nimmer worden herwonnen! Door het quantiseren gaan dus gegevens verloren en wordt de kwaliteit van het originele beeld aangetast. Het zal duidelijk zijn dat de compressiefactor afhankelijk is van de mate van

17.2 De MPEG-1 compressie van beeld en geluid

quantisering. Neem als voorbeeld de DCT-matrix van figuur 5/17.2-20.
Als gequantiseerd wordt door alle DCT-waarden te delen door 4, dan blijven er na de deling nog maar tien waarden over die niet gelijk zijn aan nul. Alle overige DCT-waarden zijn namelijk kleiner dan vier en worden tot nul gereduceerd als zij door vier gedeeld worden. Quantiseert men echter door de matrix te delen door 8, dan blijven er slechts zes waarden over die niet gelijk zijn aan nul. Het aantal nullen in de matrix neemt toe en daarmee de compressiefactor (zie verder), maar bij het herwinnen van de gegevens zal de DCT-matrix dus veel minder nauwkeurig gereconstrueerd worden dan bij quantisering door te delen door vier.

Y Quantisierungsmatrizen

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	58	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

U,V Quantisierungsmatrizen

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

Figuur 5/17.2-22: De "psychovisuele quantiseringsmatrixen" voor de luminantie- en chrominantiegegevens.

In het gegeven voorbeeld werd uitgegaan van een quantisering door middel van een constante factor 4 of 8. Bij JPEG doet men dat niet, maar wordt uitgegaan van twee zogenoemde "psychovisuele quantiseringsmatrixen". Deze zijn voorgesteld in figuur 5/17.2-22 en definiëren voor iedere frequentie in de DCT-matrixen van de luminantie en chrominantie een bepaalde unieke quantiseringsfactor. Een voorbeeld. De quantiseringsfactor voor de DC-component van de luminantie is gelijk aan 16. Dat betekent dat de DCT-factor van de DC-component wordt gedeeld door 16. Bij het voorbeeld van figuur 5/17.2-20 levert dit dus een gequantiseerde waarde van 1.260 : 16 = 79 op. Na de quantisering van de Y-waarden van de DCT-matrix van figuur 5/17.2-20 ontstaat de gequantiseerde matrix van figuur 5/17.2-23.

79	0	-1	0	0	0	0	0
-2	-1	0	0	0	0	0	0
-1	-1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Figuur 5/17.2-23: De DCT-matrix van figuur 5/17.2-20 na quantisering.

Er zijn nu nog maar zes DCT-waarden die niet gelijk zijn aan nul! Blijft de vraag hoe men aan die vreemde quantiseringsfactoren van figuur 5/17.2-22 komt. Welnu, uitgebreide onderzoeken naar de gevoeligheid van het menselijke oog voor vereenvoudiging van de luminantie- en chrominantie-gegevens in een klein beeldje hebben deze waarden als de meest gunstige opgeleverd. Ofwel, quantiseert men met deze factoren, dan zal het herwonnen beeld op het oog het minst afwijken van het niet-gequantiseerde beeld. Men houdt dus rekening met de typische (on)gevoeligheid van het oog, vandaar dat men spreekt van "psychovisuele" quantisering. Overigens is niemand verplicht zich aan deze quantiseringsfactoren te houden. Zoals reeds enige malen geschreven, MPEG-1 is een zeer open standaard, waar iedereen die daar behoefte aan heeft eigen quantiseringsfactoren kan verzinnen. Wel moeten deze factoren dan natuurlijk mee met de datastroom verstuurd worden, zodat de elektronica van de decoder deze in een register kan inlezen en gebruiken om de gequantiseerde gegevens terug om te zetten in niet-gequantiseerde DCT-waarden. Bij dat herwinnen van gegevens heeft men te maken met een tweede afwijking. Het zal duidelijk zijn dat als men een DCT-waarde deelt door de quantiseringsfactor er zelden een resultaat zonder rest ontstaat. Deze rest wordt verwaarloosd, omdat steeds wordt afgerond naar gehele getallen. Als men nadien in de decoder de gequantiseerde waarden weer vermenigvuldigd met de quantiseringsfactoren, ontstaan in de meeste gevallen andere DCT-waarden dan de originele. In het voorbeeld van figuur 5/17.2-20 levert het quantiseren van de DC-component de waarde 79 op. Gaat de decoder dit getal weer vermenigvuldigen met de quantiseringsfactor 16, dan ontstaat een herwonnen DCT-waarde van 1.264 in plaats van de originele 1.260. Ook deze afrondingsfouten dragen bij aan het lossy karakter van JPEG. Toch lijkt, ondanks al die vereenvoudigingen, weglatingen en fouten, het herwonnen beeld opmerkelijk goed op het oorspronkelijke beeld! Niet alleen een gevolg van de grote ongevoeligheid van het oog, maar ook van het feit dat DCT-waarden worden verwerkt, die niet pixelgewijs de inhoud van een macroblok bepalen, maar de algemene amplitudeverdeling van luminantie en chrominantie in dat blok. Kleine afrondingsfouten beïnvloeden dan de "heuvelachtigheid" van het amplitudeverloop van luminantie en chrominantie en hebben dan veel minder zichtbaar effect op het beeld. Het invoeren van de DCT was een gouden greep!

VOLGORDE	DECIMALE CODE	RUN LENGTH CODE
001	005	XXXX 1010
002	000	
003	000	
004	000	
005	000	
006	000	
007	000	
008	000	
009	000	
010	000	
011	006	1001 0110

Figuur 5/17.2-24: Een voorbeeld van de "run length codering".

Het vormen van de seriële datastroom

Na het quantiseren van de luminantie- en chrominantiematrixen van alle macroblokken uit de I-frames worden deze gegevens weer in de structuur van figuur 5/17.2-13 opgenomen. De gegevens worden nu zigzaggend uit de matrixen van de

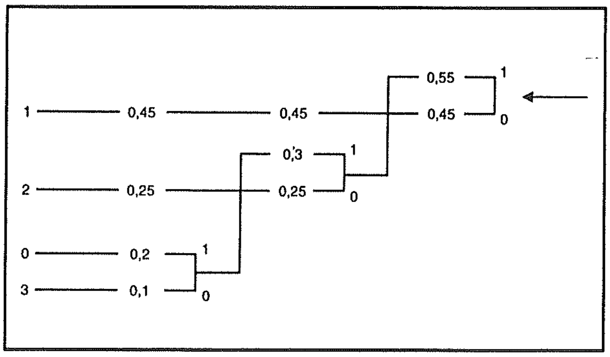
17.2 De MPEG-1 compressie van beeld en geluid

macroblokken uitgelezen, zodat een seriële datastroom ontstaat. Het gevolg van deze zigzaggende uitlezing is dat er in de code vaak een heleboel nullen achter elkaar staan.
Kijk maar naar figuur 5/17.2-22. Als de gegevens zigzaggend worden uitgelezen bestaat de seriële code van dit macroblok uit negen getallen die niet gelijk zijn aan nul en nadien uit 55 nullen. Met dergelijke gegevens kunnen compressie-algoritmen wel overweg!

De "run length codering"
Een eerste stap bij het reduceren van al deze nullen is de zogenoemde "run length codering". Het principe is vrij eenvoudig en al tientallen jaren in gebruik. Een data wordt voorgesteld door acht bits, onder de vorm van RRRR-SSSS.
In de vier R-bits wordt het aantal data gecodeerd dat gelijk is aan nul en voor de te coderen data staat. In de S-bits wordt de waarde van de data binair gecodeerd. In figuur 5/17.2-24 is een voorbeeldje ter verduidelijking gegeven. De "volgorde" geeft de seriële datastroom weer, de "decimale code" de waarde van een data uit de stroom. Data 001 heeft een waarde van 5. De codering van de R-bits is in dit voorbeeld onduidelijk omdat niet bekend is hoeveel nul-data vooraf gaan. De S-bits worden gecodeerd met "1010", de binaire representatie van het decimale getal "005". Nadien volgen 9 data die gelijk zijn aan nul. Deze worden niet gecodeerd. Het elfde data heeft "006" als decimale code. Omdat er negen "000"-data vooraf gaan, worden de R-bits gecodeerd met "1001", de binaire code voor 9. De S-bits worden gecodeerd als "0110", de binaire voorstelling van de decimale code "006". Dank zij deze codering worden de lange "000"-stromen heel efficiënt gecodeerd is slechts één byte, hetgeen een zeer grote datareductie oplevert.
De run length codering kent één uitzondering op de regel. Als na een data dat niet nul is alleen nog nul-data volgen, wordt het eerste nul-data gecodeerd met "0000-0000". Omdat bekend is dat een macroblok uit 64 data bestaat kan de decoder uit deze ene code gemakkelijk afleiden dat de rest van de data van het macroblok alleen uit nul-data bestaat. De code "0000-0000" wordt "EOB" genoemd, de afkorting van "End Of Block".

De entropie-codering via Huffmann
In de laatste stap worden de run length bytes nog eens onderworpen aan een zogenoemde "entropie-codering" volgens het Huffmann-algoritme. Huffmann was een Duitse wiskundestudent die als afstudeerproject de opdracht kreeg het meest efficiënte algoritme te ontwerpen om binaire gegevens te coderen. Dat deze man daarin meer dan geslaagd is mag blijken uit het feit dat het begrip "Huffmann-codering" een standaard is geworden in de datacommunicatie.
Bij de Huffmann-codering wordt een seriële data-stroom onderzocht op het statistisch voorkomen van bepaalde codes. Codes die het vaakst aanwezig zijn worden nadien gecodeerd met slechts één bit. Codes die iets minder vaak aanwezig zijn worden gecodeerd met twee bits. De code die het minst aanwezig is wordt gecodeerd met de langste binaire code die het algoritme toelaat. Een en ander wordt toegelicht aan de hand van figuur 5/17.2-25.
De cijfers 1 tot en met 3 staan voor codes. De hardware onderzoekt nu de procentuele aanwezigheid van deze codes in de datastroom. Code 1 is 45 % aanwezig, code 3 slechts 10 % (tweede kolom). De twee codes die het minst vaak voorkomen

krijgen nu de bits "0" en "1" toegewezen en worden nadien gecombineerd.
Hun gezamenlijke aanwezigheid bedraagt 30 %, waardoor zij opschuiven naar de tweede plaats (derde kolom). Nu wordt aan de twee codes die het minst vaak voorkomen weer een "0" en "1" toegekend, nadien worden de codes weer verenigd, hetgeen in het getekend voorbeeld een voorkomen van 55 % oplevert. Deze code komt nu dus het vaakst voor, zodat zij oprukt naar de eerste plaats (vierde kolom). Ook nu wordt weer "0" en "1" toegekend. In dit eenvoudige voorbeeld, waar slechts vier mogelijke codes worden onderzocht, beëindigt het Huffmann-algoritme het voorbereidend werk. De omcodering wordt in een matrix opgenomen en de datastroom wordt een tweede maal doorlopen, waarbij de omcodering tot stand komt.



Figuur 5/17.2-25: Het Huffmann-algoritme verklaard aan de hand van een eenvoudig voorbeeld.

Komt de hardware code "1" tegen, dan wordt deze gecodeerd door een enkelvoudig bit "0". Daarbij maakt het niet uit hoe lang code 1 is! Komt het systeem code 2 tegen, dan wordt dit omgezet in de bitvolgorde "1-0". In het getekende voorbeeld wordt code 3 omgezet in de langste bitcombinatie, namelijk "1-1-0". De Huffmann-codes worden dus van rechts naar links samengesteld door de boomstructuur te volgen van het rechter eindpunt tot aan de betreffende code. Het zal duidelijk zijn dat ook dit systeem een grote datareductie tot gevolg heeft. De RRRR-SSSS codes van de run length codering, ieder goed voor acht bit, worden via Huffmann omgezet in codes, die gemiddeld genomen veel korter zijn. In figuur 5/17.2-26 is hiervan een voorbeeldje getekend.

RRRR SSSS	Huffmann-Code-Word
0000 0000	1010 (EOB)
0000 0001	00
0000 0010	01
0000 0011	100
0000 0100	1011
0000 0101	11010
0000 0110	1111000
0000 0111	11111000
0000 1000	1111110110
0000 1001	1111111110000010
0000 1010	1111111110000011

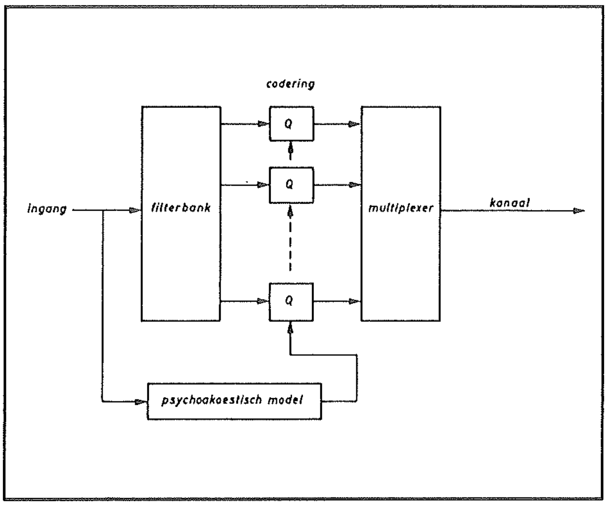
Figuur 5/17.2-26: Het omzetten van de run length codes in Huffmann-codes.

De MPEG-1 codering van audio

Inleiding

MPEG-1 gebruikt voor het coderen van audio een algoritme dat in basis niet afwijkt van de PASC-codering die Philips

gebruikt bij het digital compact cassette (DCC) systeem. Er wordt dus gebruik gemaakt van "perceptuele subbandcodering". Het volledige audiospectrum wordt ingedeeld in 32 subbanden, op ieder band worden maskeringstechnieken toegepast om geluidsinformatie, die toch niet door het menselijk oor wordt opgevangen, te verwijderen. Het blokschema van de audiocompressie is getekend in figuur 5/17.2-27.



Figuur 5/17.2-27: De audiocodering volgens MPEG-1.

Lagenconcept

Ook de audiocodering van MPEG-1 is een vrij in te vullen standaard. De standaard schrijft de algemene principes voor, de gebruiker heeft een enorme vrijheid om eigen compressie-algoritmen te definiëren en toe te passen. De standaard geeft ruimte aan drie zogenoemde "layers" of lagen. Alle lagen kennen bemonsteringsfrequenties van 32 kHz, 44,1 kHz en 48 kHz en bieden vier werktoestanden:

- monofoon geluid;
- tweekanaals geluid;
- stereofoon geluid;
- joint stereo.

De laatste werktoestand zoekt de zogenoemde "redundantie" oftewel identieke gegevens in beide stereokanalen op en verzendt deze informatie slechts één maal.
Dit biedt ruimte in de op Video-CD beschikbare datastroom van ongeveer 200 kbit/s om het geluid met een hogere kwaliteit te versturen.

Layer 1

Laag 1 van MPEG-1 is het eenvoudigst opgezet. De specificaties voorzien in filtering naar 32 subbanden van gelijke bandbreedte, adaptieve bittoewijzing en blokcompressie op frames van 8 ms breedte. De resulterende datastroom kan ingesteld worden tussen 32 kbit/s tot 448 kbit/s door het aanpassen van de compressie-algoritmen.
Layer 1 werd voornamelijk gedefinieerd voor goedkope consumententoepassingen, waarbij de noodzakelijke decodeerelektronica in één IC geïntegreerd kan worden.

Layer 2

Deze laag wordt ook wel "MUSICAM' genoemd. Deze laag werkt met speciale compressietechnieken, waardoor de gegevensreductie toeneemt. Er wordt namelijk gecomprimeerd over veel langere frames. De noodzakelijke decodeerelektronica is echter ingewikkelder en dus duurder.

Layer 3

Deze laag biedt zeer goede specificaties, dank zij zeer uitgebreide en ingewikkelde compressie-algoritmen. Bij deze laag wordt bijvoorbeeld gebruik gemaakt van Huffmann-codering. Het gevolg is dat deze laag zelfs bij een monofone datastroom van slechts 64 kbit/s een zeer hoge geluidskwaliteit oplevert.

17.2 De MPEG-1 compressie van beeld en geluid

Geïntegreerde dynamiekcompressie

Zoals bekend wordt het meeste geluidsmaterieel begrensd in dynamisch bereik. Het dynamisch bereik geeft de verhouding tussen het "stilste" geluid en het "hardste" geluid in dB. Het toepassen van compressie is noodzakelijk om er zeker van te zijn dat de geluidsverwerkende apparatuur nooit overstuurd wordt. De audionormering van MPEG-1 biedt de mogelijkheid de compressiefactoren die in de studio worden bepaald mee te verzenden met de datastroom, zodat aan de ontvangerzijde de mogelijkheid bestaat de originele dynamiek van het geluid weer op een heel eenvoudige manier te herstellen.

5/17.3

Het formaat van Video-CD

Inleiding

Drie lagen
Bij het bespreken van het formaat waarin de gecomprimeerde video- en audiogegevens weggeschreven worden naar de spiraal van de Video-CD moet men onderscheid maken in drie verschillende organisatorische lagen:
- sectorstructuur;
- logische structuur;
- gegevensstructuur.

De sectorstructuur
De eerste laag beschrijft de sectorstructuur van de Video-CD. Dit is de manier waarop de spiraal is ingedeeld in sectoren en hoe de nuttige gegevens, samen met sync-, header-, subheader-, EDC-, ECC- en controle-bytes in deze sectoren zijn opgenomen.

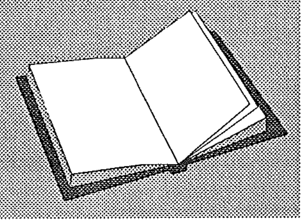
De logische structuur
De logische structuur beschrijft de manier waarop de gegevens hiërarchisch op de Video-CD staan, met andere woorden de indeling van de Video-CD in directories en subdirectories.

De gegevensstructuur
De gegevensstructuur beschrijft hoe de eigenlijke video- en audiogegevens gecombineerd worden tot een seriële datastroom die uiteindelijk in de spiraal in sectoren wordt geschreven.

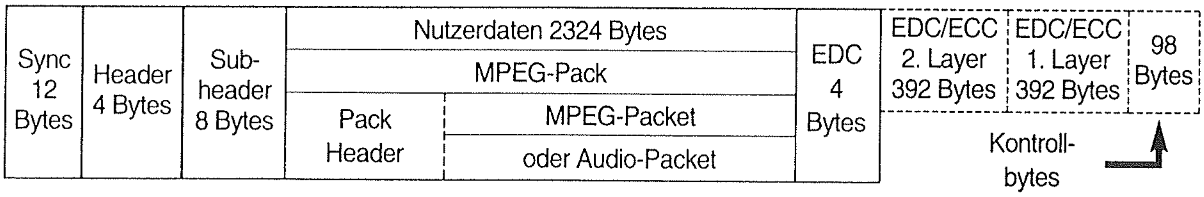
De sectorstructuur

ISO 9660
De sectorindeling van een Video-CD voldoet aan de ISO 9660 specificaties. Dat betekent dat een dergelijk schijfje zonder problemen uit te lezen is op moderne CD-ROM drives die aan deze specificaties voldoen. Omgezet naar CD-nomenclatuur kan men stellen dat Video-CD's beschreven zijn volgens de CD-ROM XA Mode-2 Form-2 standaard. Dit is hetzelfde formaat dat ook wordt toegepast bij Photo-CD's. Vandaar dat alle apparatuur die in staat is Photo-CD's te verwerken ook in staat is Video-CD's af te spelen. Er bestaat echter toch een belangrijk verschil!

LEES OOK:
Hoofdstuk 5/8.1
Hoofdstuk 5/16.2
Hoofdstuk 5/17.2
Hoofdstuk 5/17.4



17.3 Het formaat van Video-CD



Figuur 5/17.3-1: De samenstelling van een sector van een Video-CD.

De plaatjes op Photo-CD's zijn volledig softwarematig te verwerken, bij Video-CD lukt dit niet. Het decoderen van de video-informatie kost softwarematig te veel tijd, zodat hiervoor extra elektronica onder de vorm van een MPEG-1 decoder en een zeer snelle processor noodzakelijk is.

De algemene structuur van een sector van een Video-CD

De algemene samenstelling van een sector van een Video-CD is getekend in figuur 5/17.3-1.
Men onderscheidt:
- de sync-bytes;
- de header;
- de sub-header;
- de gegevens;
- de EDC/ECC-bytes;
- de control-bytes.

Een korte beschrijving van de functie van deze gegevens.

De sync-bytes

Een sector start met 12 synchronisatiebytes. Om het opzoeken van de startsector van een video- of audioframe te versnellen worden de sectoren bij het XA Form-2 formaat duidelijk fysisch van elkaar gescheiden door het tussenvoegen van 12 synchronisatiebytes. Deze bevatten een unieke code, die gemakkelijk door de elektronica van de Video-CD speler te detecteren is.

De header

De vier bytes van de header dienen voor het adresseren van de sector met een tijdcodering in minuten, seconden en sectoren.
De inhoud van deze vier bytes is als volgt:
- byte 1: minuten;
- byte 2: seconden;
- byte 3: sectoren;
- byte 4: mode.

Dit laatste byte heeft niets met de tijdcodering te maken, maar definieert de mode van de sector:
- Mode-0: geeft aan dat de sector leeg is;
- Mode-1: de sector is ingedeeld volgens het XA Form-1 formaat;
- Mode-2: de sector is ingedeeld volgens het XA Mode-2 formaat.

De sub-header

De acht bytes in de sub-header identificeren het soort gegevens dat in de sector is geschreven. Deze gegevens ondersteunen het zogenoemde ***"interleaving principe"***, waarmee het mogelijk is video- en audiogegevens door elkaar te husselen, zodat de gegevens in de meest voor de hand liggende volgorde van de Video-CD gelezen kunnen worden.

17.3 Het formaat van Video-CD

- **Bytes 1 en 5**
  In de bytes 1 en 5 van de sub-header wordt het bestandsnummer geschreven. Dat nummer heeft een waarde van 0 tot en met 255. De waarde 0 wil zeggen dat de sector geen deel uitmaakt van het interleaved schema en dat alle opeenvolgende sectoren met code 0 achter elkaar als continue datastroom uitgelezen kunnen worden.
- **Bytes 2 en 6**
  In bytes 2 en 6 wordt een zogenoemd kanaalnummer opgeslagen, ook een getal tussen 0 en 255. Het kanaalnummer wordt gebruikt om het uitlezen van interleaved sectoren te vergemakkelijken. Video-data worden gekarakteriseerd door kanaalnummers tot en met 31.
- **Bytes 3 en 7**
  Bytes 3 en 7 noemt men de "sub-mode"-bytes. De acht bits zijn vlaggen die op logisch "0" of "1" gezet kunnen worden en de volgende gegevens bevatten:
  - Bit 0:
    Wordt hoog gezet als de sector de laatste van een record is.
  - Bit 1:
    Wordt hoog gezet als de sector video-informatie bevat.
  - Bit 2:
    Wordt hoog gezet als de sector ADPCM-audio informatie bevat.
  - Bit 3:
    Wordt hoog gezet als de sector coderingsdata bevat.
  - Bit 4:
    Zet de "trigger" aan of uit, een signaal dat voor bepaalde elektronische bewerkingen in de XA-controller noodzakelijk is.
  - Bit 5:
    Bepaalt het Form-1 formaat of Form-2 formaat van de sector.
  - Bit 6:
    Wordt hoog gezet als de sector zogenoemde "real-time" gegevens bevat. Real-time gegevens zijn gegevens die zo snel mogelijk vanuit de optische spiraal op de Video-CD naar het geheugen van de speler overgebracht moeten worden. Wordt bit 6 hoog gezet, dan worden de foutdedectietechnieken voor de sector uitgeschakeld.
  - Bit 7:
    Wordt hoog gezet als de sector de laatste sector van een bestand is, dit is het EOF-bit.
- **Bytes 4 en 8**
  De bytes 4 en 8 bevatten informatie over de manier waarop audio- en videogegevens zijn gecodeerd. Een bit bepaalt bijvoorbeeld of de audiogegevens mono of stereo zijn. Een tweede bit bepaalt de zogenoemde "pre-emphasis" van het geluid. Staat dit bit op "1", dan wordt de audioprocessor opgedragen alle hoge tonen met 6 dB te verzwakken, zodat de ruis in het geluid afneemt. De overige bits definiëren de compressietechnieken, die op de gegevens in de sector zijn toegepast.

De gegevens
Nadien staan 2.324 bytes ter beschikking voor het schrijven van de gegevens. Hierin wordt ofwel een MPEG-Pack geschreven, ofwel een Pack Header met het daaropvolgende MPEG-Packet. Deze begrippen worden later uitgelegd.

De EDC- en ECC-bytes
Deze bytes worden gebruikt voor de foutcontrole en -correctie van de gegevens. Uiteraard wordt ook bij Video-CD gebruik gemaakt van de ***"Cross Interleaved Reed-Solomon Code"***, afgekort tot "CIRC".

De controle-bytes

Een sector wordt afgesloten met de 98 controle-bytes, die ook van de Audio-CD bekend zijn en beschreven staan in de "Red Book"-specificaties. Niet duidelijk is of deze bytes gebruikt worden.

De indeling in track's

De sectoren worden ingedeeld in zogenoemde "track's". Track's bevatten sectoren met identieke informatie, bijvoorbeeld video- of audio-informatie. De modus mag in een track zelfs niet eens veranderen. Men zegt wel eens dat "de kleur van een track niet mag veranderen". Dat betekent dat alle sectoren in een en dezelfde track volgens dezelfde "Book"-specifikaties geschreven moeten zijn. De track's zijn via de informatie in de TOC, de "Table Of Contents" toegankelijk. Dat betekent dat het mogelijk is de video-informatie die in een bepaalde track start rechtstreeks te benaderen. Video-CD biedt dus de mogelijkheid niet alleen een lange speelfilm weer te geven, maar ook talrijke kleine videosequenties, die ieder in een track staan. Iedere track begint met de "Pre-gap", een ongeveer twee seconde durende digitale pause. In dit gebied worden alleen nullen geschreven. In de eerste track van de Video-CD wordt het besturingsprogramma ondergebracht, de volgens ISO 9660 georganiseerde gegevensstructuur van de CD en optioneel Karaoke-informatie. Vanaf de tweede track volgt dan de video- en audio-informatie.

De logische structuur

Directories

Als gevolg van de ISO 9660 specificatie kan men een Video-CD beschouwen als een diskette. Dat betekent dat alle informatie in directories is opgenomen. De "root" van een Video-CD bevat in ieder geval de volgende directories:

- CDI;
- MPEGAV;
- VCD.

Daarnaast kan er nog een directory KARAOKE aanwezig zijn.

CDI

Deze directory bevat de programma's en gegevensbestanden die noodzakelijk zijn om de Video-CD in een CD-I speler af te spelen.

MPEGAV

Deze directory bevat verwijzingen waarmee het systeem kan gestuurd worden naar de individuele track's.

VCD

Deze directory bevat in ieder geval twee bestanden:

- INFO.VCD
  Dit bestand bevat de systeemidentificatie van de Video-CD en de identifikatie van de individuele schijf.
- ENTRIES.VCD
  Dit bestand koppelt de zogenoemde "Entry Points" van alle sequenties op de Video-CD aan de namen van de sequenties. Entry Points zijn track's die onmiddellijk, via de software, toegankelijk zijn en waarvan men dus de inhoud direct op het beeldscherm kan weergeven. Dank zij dit bestand kan een Video-CD bijvoorbeeld een menuutje op het scherm zetten, waardoor het mogelijk is met de muis of de afstandsbediening bepaalde selecties uit te voeren. Iedere sequentie wordt in een van maximaal 98 mogelijke track's gestart. Iedere sequentie kan daarnaast

nog eens verwijzingen hebben naar maximaal 98 direct toegankelijke subtrack's van de betreffende sequentie. In totaal lijkt het alsof een Video-CD dus maximaal 9.604 direct via de software toegankelijke Entry Points heeft! Dat is niet zo, want het maximaal aantal wordt beperkt tot 500. Nog heel veel en natuurlijk zal niet iedere softwareleverancier van al deze mogelijkheden gebruik maken. Zo lijkt het voor het afspelen van een speelfilm weinig zinvol zoveel startposities in te voeren. Maar dit uitgebreide zoeksysteem is natuurlijk voor tal van andere toepassingen uitermate interessant. Zo is het, via nog te ontwikkelen software, zonder meer mogelijk dat men bepaalde videosequenties van een Video-CD naar een (niet gefragmenteerd) gedeelte van een harde schijf kopieert en nadien deze MPEG-1 sequentie als normaal bestand kan gebruiken in Multimedia-presentaties, als ware het een AVI-, WAV- of MIDI-bestand!

KARAOKE

In deze directory bevinden zich bestanden die noodzakelijk zijn om videosequenties te koppelen aan de diverse audiosequenties die in diverse talen ter beschikking staan. Bovendien bevat deze directory verschillende karakterset's voor de ondersteunde talen, zoals bijvoorbeeld het Kanji voor het Japans.

De gegevensstructuur

Inleiding

De gegevensstructuur beschrijft hoe de met MPEG-1 gecodeerde en gecomprimeerde seriële datastromen van audio en video worden gecombineerd. Deze afzonderlijke datastromen worden verpakt in zogenoemde "Packet's". Dat deze stromen gecombineerd moeten worden zal duidelijk zijn. Het is immers de bedoeling dat speelfilms lipsynchroon worden afgespeeld en dat kan alleen maar als het systeem met niet al te grote vertragingen video- en audiogegevens van de CD-spiraal kan lezen. De Packet's worden verenigd tot grotere datastructuren die "Pack's" worden genoemd. Eén Pack kan zowel video als audio Packet's bevatten. Uit de in het "Red Book" vastgelegde specificaties van de CD volgt dat er maximaal 75 sectoren per seconde worden uitgelezen. Dat betekent dat per seconde maximaal 75 x 2.324 = 174,3 kB kan worden gelezen. Vanwege de beperkingen van de MPEG-compressie moet hiervan 144 kB gereserveerd worden voor video zodat ongeveer 28 kB ter beschikking staat voor het geluid. Om de synchronisatie tussen beeld en geluid te bewaren moet er ongeveer één audio Packet per vijf video Packet's verstuurd worden.
Het algemeen schema van de gegevensstructuur is geschetst in figuur 5/17.3-2.

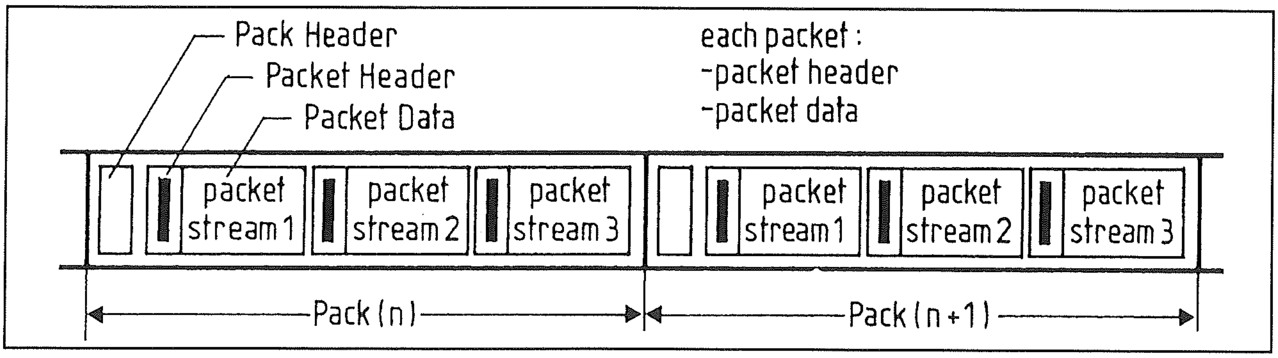
Headers

Zowel de Packet's als de Pack's worden voorafgegaan door Headers. In deze Headers wordt beschreven welke data in de Packet's en de Pack's zijn opgenomen en hoe die gegevens gecodeerd zijn.

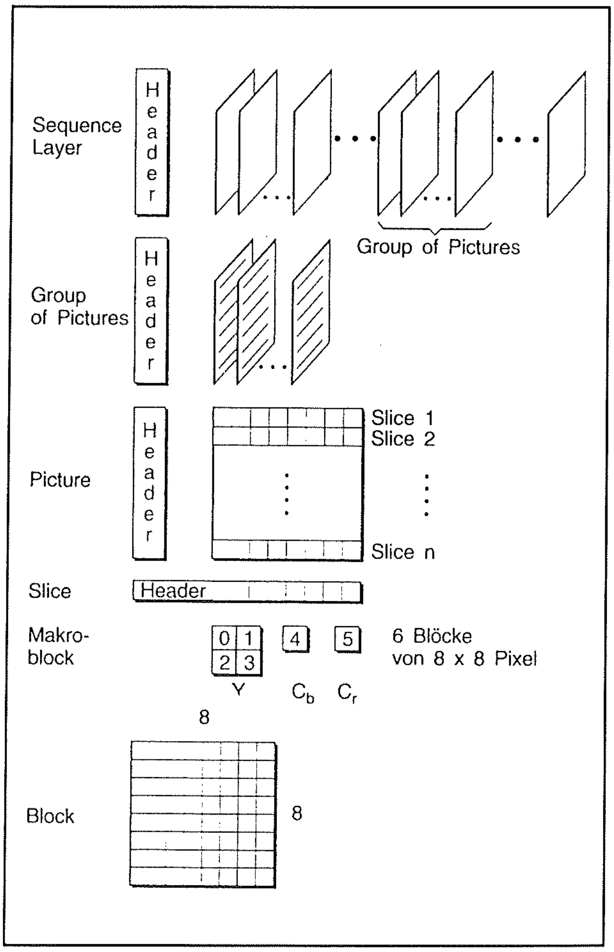
De datacodering van MPEG-1 video

De videogegevens worden op een vrij ingewikkelde manier gecombineerd tot een seriële datastroom. Het volledige schema is weergegeven in figuur 5/17.3-3, waarbij de complexiteit van de codering oploopt van het eenvoudigste onderste niveau tot het uiteindelijke bovenste niveau.

17.3 Het formaat van Video-CD



Figuur 5/17.3-2: De algemene gegevensstructuur van de seriële gecombineerde datastroom.



Figuur 5/17.3-3: De codering van de videogegevens.

Block's

Het basismateriaal zijn de Block's van 8 bij 8 pixels. Deze bevatten de MPEG-1 gecodeerde data van de chrominantie en de luminantie gegevens, dus de Y-, C_b- en C_r-signalen. Nadien worden deze Block's samengevat tot zogenoemde Macro-block's, waarbij om in hoofdstuk 5/17.2 beschreven redenen steeds vier Y-Block's gecombineerd worden met twee Block's die de chrominantiegegevens bevatten. Ieder Macro-block wordt voorafgegaan door een Header, waarin de positie van het Macro-block in het totale beeld beschreven wordt.

De Slices

In de zogenoemde "Slices" worden naast elkaar gelegen Macro-block's ondergebracht. Deze hoeven niet aan de rand van het beeld te beginnen, maar kunnen net zo goed ergens midden in een beeldlijn liggen.
De MPEG-1 decoder kan de gegevens in een Slice fundamenteel decoderen zonder dat hiervoor een beroep moet worden gedaan op gegevens in andere Slices. Op deze manier wordt verhinderd dat fout gelezen gegevens zich door het beeld gaan voortplanten. Een fout gelezen bit, dat door de correctie komt, kan nu alleen die beeldelementen aantasten die tot één Slice behoren. De Header die voorafgaat aan de Slice bepaalt uiteraard de plaats van de gegevens in het beeld.

17.3 Het formaat van Video-CD

GOP's

Slices worden gecombineerd tot een "Picture". Hierin worden alle gegevens opgenomen die noodzakelijk zijn om een volledig frame te decoderen. Een "Picture" wordt weer voorafgegaan door een Header, waarin voornamelijk informatie staat over het soort frame dat wordt ingelezen, met name of het Picture een I-, P- of B-frame bevat. De MPEG-1 decoder moet namelijk heel andere algoritmen toepassen voor deze drie fundamenteel van elkaar verschillende beelden.
De afzonderlijke frames worden samengevat in zogenoemde "GOP's", afkorting van "Group Of Pictures". Deze GOP's kunnen een willekeurig aantal frames bevatten, maar wel is het noodzakelijk dat in iedere GOP minstens een I-frame aanwezig is. Zonder deze volledige beschrijving van een beeld kan de MPEG-1 decoder niets aanvangen met de gegevens in de P- en B-frames, die immers vanuit een voorafgaand I-frame worden berekend. In de GOP-Header worden gegevens vermeld over de aard van de beelden in de groep, maar tevens informatie over:
- horizontale en vertikale resolutie;
- beeldfrequentie;
- quantiseringsmatrixen.

De decoder heeft deze gegevens nodig, omdat het dank zij de vele Entry Points kan voorkomen dat de decoder geconfronteerd wordt met een GOP waarin de gegevens niet standaard zijn gecodeerd. Het is zelfs mogelijk voor iedere GOP geheel eigen quantiseringsmatrixen te definiëren!

Sequence Layer

De GOP's worden weer samengevat in zogenoemde "Sequence Layer's". Dat zijn de uiteindelijke pakketten videogegevens, die in de video-Packet's terecht komen.

17.3 Het formaat van Video-CD

5/17.4

De MPEG-1 coders en decoders

Inleiding

Markt in beweging

Het zal duidelijk zijn dat alle grote IC-fabrikanten bezig zijn met het ontwikkelen van MPEG coders en decoders. De toepassingen voor deze schakelingen zijn legio, van gebruik in CD-I spelers met FMV-ondersteuning tot in digitale beeldtelefoons. Vanwege de nieuwe technologie is de markt natuurlijk volop in beweging. Was het tot voor kort nog zo, dat voor een volledige MPEG-1 decoder een handjevol IC's noodzakelijk was, op dit moment leveren reeds verschillende halfgeleiderfabrikanten een één-chip oplossing, waar alleen een extern geheugen en een besturingsprocessor aan toegevoegd moeten worden. Het is niet de bedoeling van dit hoofdstuk alle MPEG IC's tot in de details te beschrijven. Wel wordt een algemeen inzicht gegevens in de blokschematische werking van deze zeer ingewikkelde schakelingen.

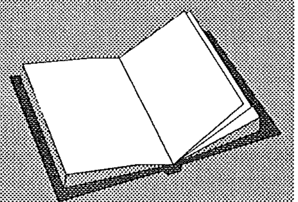
Marktoverzicht

Op het moment waarop dit hoofdstuk geschreven wordt (december 1994) is bekend dat onderstaande fabrikanten ofwel reeds MPEG-chips's in de aanbieding hebben, ofwel druk bezig zijn met het ontwikkelen van dergelijke schakelingen. Helaas valt het hierbij op dat deze toekomstgerichte technologie bijna volledig in handen is van Amerikaanse bedrijven. Alleen het Franse SGS-Thomson speelt dank zij haar *"Image Processing Business Unit"*-laboratorium in Grenoble een toontje mee!

- Array Microsystems
  Dit bedrijf levert de A77100 MPEG-1 videocoder en de A77300 MPEG-1 videodecoder.
- AT&T Microelectronics
  Met de slechts $ 50,00 kostende AV6101 kan men zowel MPEG-1 als MPEG-2 coderen.
- C-Cube Microsystems
  Dit bedrijf is een van de pioniers op het gebied van MPEG-chip's en was minstens één jaar lang de enige leverancier van bruikbare schakelingen. C-Cube levert onder de verzamelcode CL-xxxx een aantal IC's die toegepast kunnen worden in MPEG-coders en -decoders. De CL-480 is een zeer moderne ontwik-

LEES OOK:

Hoofdstuk 5/17.1
Hoofdstuk 5/17.2



keling. Dit IC bevat zowel een MPEG-1 video- als audiodecoder en bovendien een zeer snelle processor uitgevoerd in RISC-technologie.

- Cirrus Logic
  Dit bedrijf levert de CL-GD7542 decoder, een energiezuinige schakeling speciaal ontwikkeld voor de toepassing in draagbare apparatuur.
- IBM Microelectronics
  Ook "Big Blue" blijft niet achter en is bezig met het ontwikkelen van een één-chip MPEG-1 en -2 systeem.
- LSI Logic
  Dit bedrijf levert de L64000 videocoder voor MPEG-1 en -2. Een identieke decoder is onder code L64002 in ontwikkeling.
- Oak Technology
  Deze bekende fabrikant van grafische kaarten voor PC's ontwikkelt de OTI-201, een volledige decoder voor video MPEG-1 en -2.
- SGS-Thomson Microelectronics
  Dit is de enige Europese fabrikant die enthousiast op de MPEG-trein gesprongen is. Dit bedrijf levert verschillende JPEG- en MPEG-decoders en -coders onder de verzamelcode STIxxxx.
- Texas Instruments
  Naast zogenoemde DSP's, digitale signaal processoren die toegepast kunnen worden in de MPEG-technologie, is dit bedrijf bezig met de ontwikkeling van de AV220 en AV120 MPEG-1 videocoders en -decoders.
- Western Digital
  Dit bedrijf is bekend van zijn grafische processoren van de WD-serie. De nieuwe ontwikkeling WD9710 gaat een stapje verder. Dit IC bevat een volledige MPEG-1 en -2 videodecoder, gecombineerd met een grafische processor met een breedte van 64 bit en acceleratiefuncties. Bovendien is een 3 x 24 bit DAC geïntegreerd, zodat het IC in staat is "true-color"-beelden met een kleurendiepte van meer dan 16 miljoen kleuren op het scherm te zetten. Dit IC luidt een nieuwe ontwikkeling in, waarbij MPEG decodeerfuncties standaard op grafische kaarten en later zelfs op moederborden voor PC's worden ondergebracht.
- Zoran Corporation
  Dit bedrijf heeft op dit moment de ZR36100 MPEG-1 videodecoder in de aanbieding.

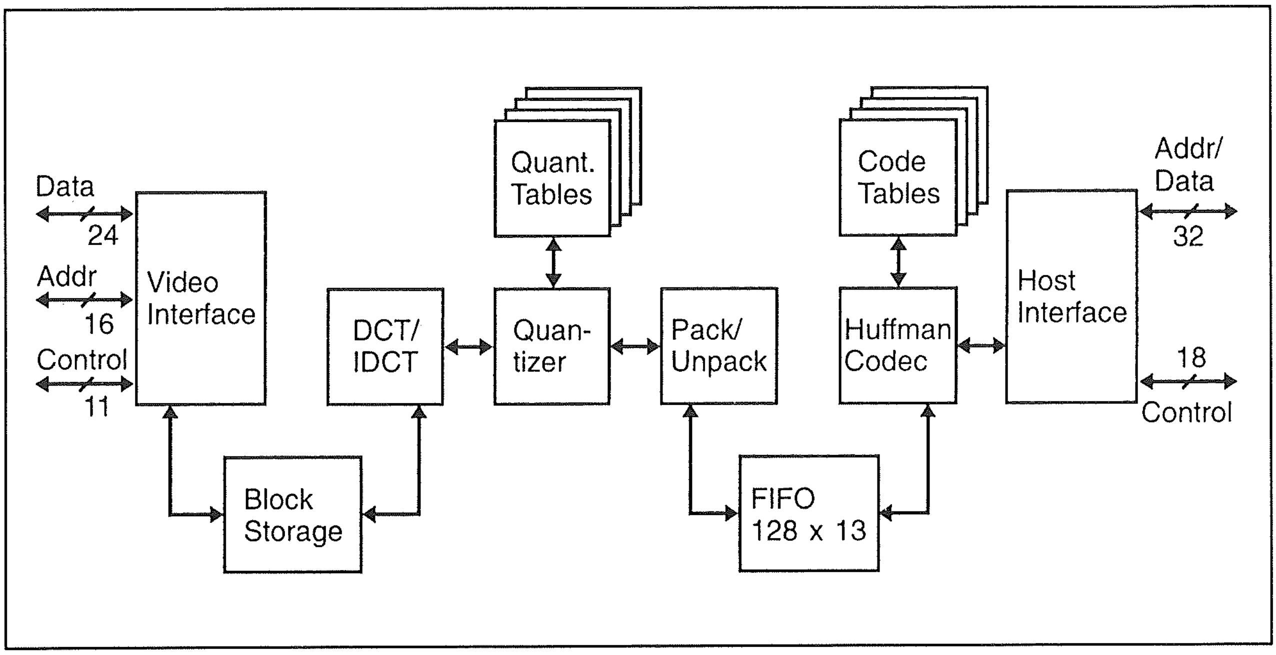
JPEG coders en decoders

Het hart van het MPEG-systeem
De eenvoudigste IC's zijn de JPEG-coders/decoders. Dergelijke schakelingen vormen uiteraard het hart van een volledig MPEG-systeem. De IC's kunnen de volgende bewerkingsstappen op het videosignaal uitvoeren:
- DCT-transformatie;
- Quantisering;
- Run Lenght Codering;
- Huffman-codering.

Bovendien werken dergelijke IC's bidirectioneel, hetgeen wil zeggen dat een en dezelfde schakeling zowel bruikbaar is voor de codering als voor de decodering.

De CL-550 als voorbeeld
Het blokschema van een dergelijke schakeling, in dit geval de CL-550 van C-Cube, is getekend in figuur 5/17.4-1.
De "Video Interface" verzorgt de koppeling en synchronisatie van het IC aan en met de ingangsgegevens. De gegevens worden tijdelijk opgeslagen in de "Block Storage".

17.4 De MPEG-1 coders en decoders



Figuur 5/17.4-1: Het blokschema van een JPEG coder/decoder.

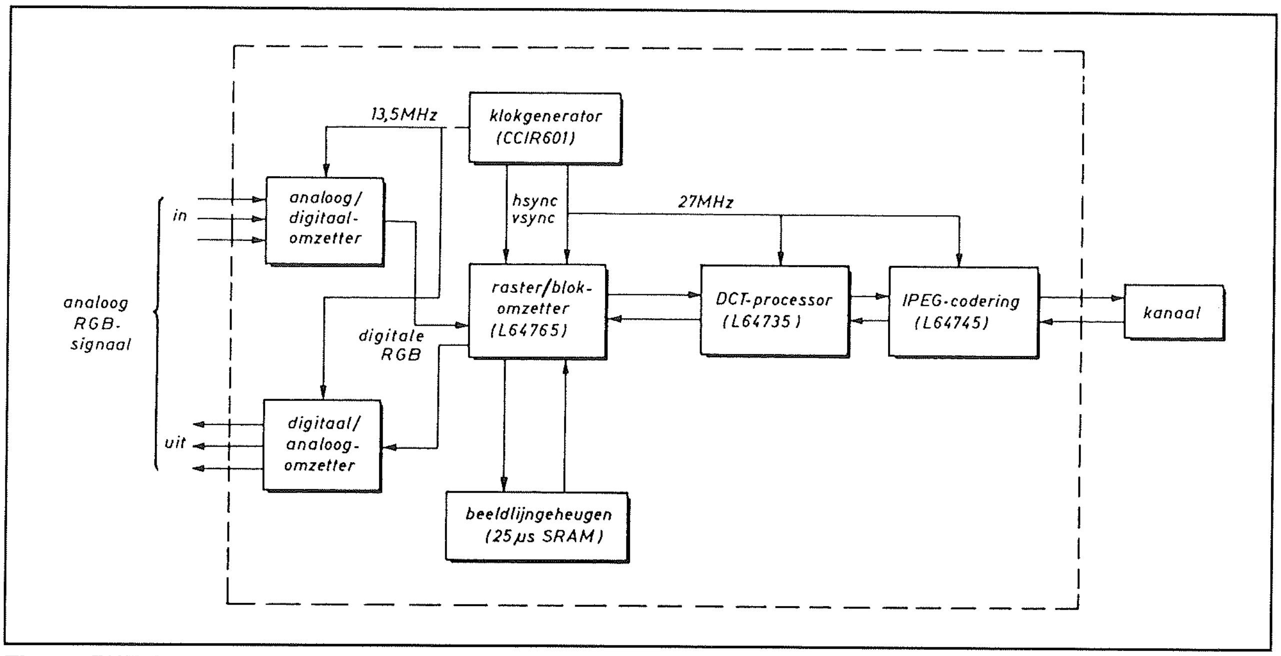
Dit is een geheugen waarin de gegevens gegroepeerd worden in Block's. De "DCT/IDCT" voert de discrete cosinus transformatie op de gegevens uit. Nadien volgt de quantisering in de "Quantisizer", waarbij natuurlijk de in een geheugen opgeslagen tabellen voor luminantie en chrominantie noodzakelijk zijn. Deze tabellen moeten extern geladen worden, waardoor de mogelijkheid bestaat voor iedere toepassing individuele tabellen samen te stellen. Het blok "Pack/Unpack" verzorgt de RLE-codering van de datastroom. Nadien volgt de "Huffman-CODEC", waarin aan de hand van een codetabel vaak voorkomende bitvolgorden worden omgezet in kortere codes.

De JPEG chipset van LSI Logic

LSI Logic levert voor de bi-directionele JPEG-bewerkingen een set van drie IC's. Het blokschema van het systeem is getekend in figuur 5/17.4-2. De set bestaat uit:
- de raster/blok omzetter L64765;
- de DCT-coprocessor L64735;
- de JPEG-coder L64745.

Het systeem moet aangevuld worden met een klokgenerator, een ADC en een DAC en twee statische lees/schrijfgeheugens. De binnenkomende gegevens bestaan uit drie videostromen in de vorm van analoog-RGB. De ADC converteert dit naar digitale signalen. De gegevens worden door de raster/blok omzetter omgezet in $Y\text{-}C_r\text{-}C_b$ en opgeslagen in een beeldlijngeheugen onder de vorm van 8 bij 8 blokken. De L64735 kan zowel de DCT als IDCT uitvoeren op de gegevens. Een volledige acht bij acht transformatie wordt in 64 clockcyclussen uitgevoerd. Vanwege de maximale clockfrequentie van 35 MHz is dit IC in staat de DCT real-time uit te rekenen. De DCT-processor kan de gegevens in zigzag vorm uitvoeren, maar ook gewoon serieel. De L64745 verzorgt de entropiecodering van de gegevensstroom. De gecomprimeerde gegevens wordt opgeslagen in een geïntegreerd FIFO-register met een breedte van 32 bit en een diepte van 32 woorden.

17.4 De MPEG-1 coders en decoders



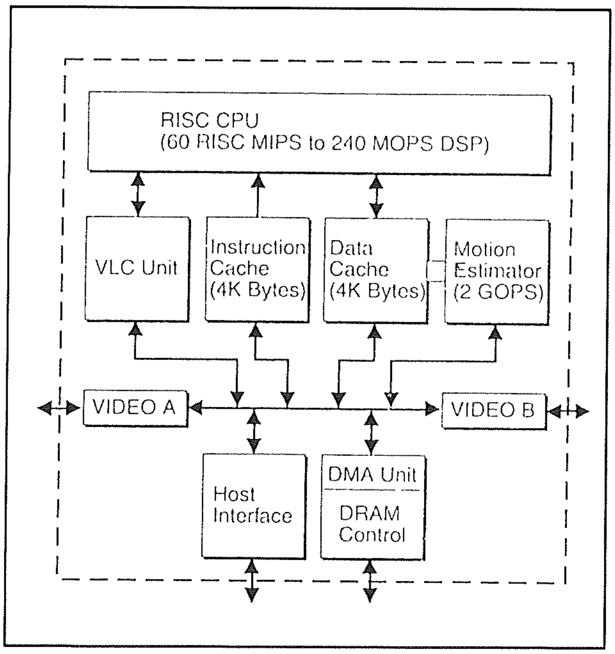
Figuur 5/17.4-2: De chipset van LSI Logic.

MPEG videocoders

Blokschema

Als typisch voorbeeld van een volledige MPEG-1 coder wordt het blokschema van de CL-4000 van C-Cube in het kort besproken. Dit IC is een zogenoemde "VRP", een "Video RISC Processor". Het blokschema van dit IC is getekend in figuur 5/17.4-3. Het IC is vrij programmeerbaar, hetgeen wil zeggen dat de CL-4000 zowel bruikbaar is voor het in real-time coderen volgens de MPEG-1, MPEG-2, JPEG en H.261 normen. Deze laatste standaard is een professionele internationale afspraak, die gebruikt wordt in digitaal werkende TV-studio's.

De VRP verwacht dat de aangeleverde videogegevens in Y-C_b-C_r formaat staan. In het blok "VIDEO A" wordt besloten of de gegevens onveranderd doorgelaten kunnen worden, of dat het de bedoeling is de resolutie om te rekenen naar de standaardnorm voor MPEG-1. Dit blok zet de gegevens bovendien in de juiste structuur op de interne bus, die bestuurd wordt door een RISC-processor met een verwerkingssnelheid van maximaal 240 MOP's.

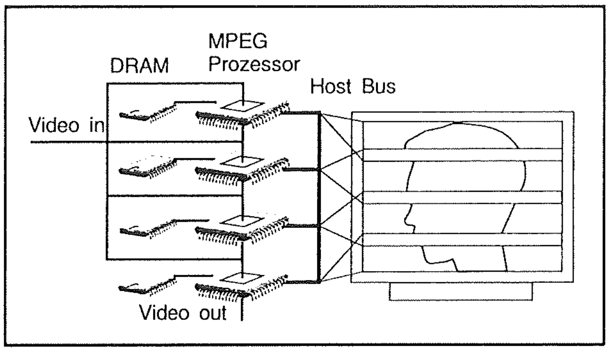


Figuur 5/17.4-3: Het blokschema van de MPEG-coder CL-4000 van C-Cube.

17.4 De MPEG-1 coders en decoders

Dat betekent dat deze processor in staat is 240 miljoen Micro Operations per Second uit te voeren. De "Motion Estimator" is in feite een tweede processor, die als coprocessor voor de RISC werkt en onder controle van het bedrijfssysteem de berekening van de gunstigste verplaatsingsvectoren en -fouten voor zijn rekening neemt. Deze processor kan maximaal 2 GOP's aan, hetgeen betekent dat de schakeling in staat is twee miljard operaties per seconde uit te voeren.
De DCI, de quantisering en de RLE- en de entropie-compressie worden allemaal volledig softwarematig door de snelle RISC-processor uitgevoerd. Op de CL-4000 moet een geheugen van het type DRAM worden aangesloten. In dit geheugen wordt uiteraard de microcode van het besturingssysteem opgeslagen, maar ook alle beeldjes die noodzakelijk zijn voor het berekenen van de P- en B-frames uit een I-frame. De omvang van dit geheugen is meestal 4 MB, een waarde die ook bij het omgekeerde proces, dus het decoderen, absoluut noodzakelijk is.

Parallelle verwerking
Bij niet zo snel werkende schakelingen moet men, als real-time een absolute eis is, overschakelen op parallelle verwerking van de gegevens. Het blokschema van een dergelijk systeem is getekend in figuur 5/17.4-4. Dit systeem is op dit moment het enige waarmee de kwalitatief hoogwaardige formaten HDTV, EDTV en SDTV van MPEG-2 in real-time verwerkt kunnen worden. De binnenkomende hoge resolutie videobeelden worden opgedeeld in horizontale stroken, die ieder door een eigen MPEG-processor met eigen DRAM verwerkt worden. Voor de noodzakelijke berekeningen van de verplaatsingsvectoren is het echter noodzakelijk dat de systemen over elkaars informatie kunnen beschikken. Het zal duidelijk zijn dat van een "eenvoudige" parallelle verwerking geen sprake is, maar dat het systeem ongelooflijk complex in elkaar zit.

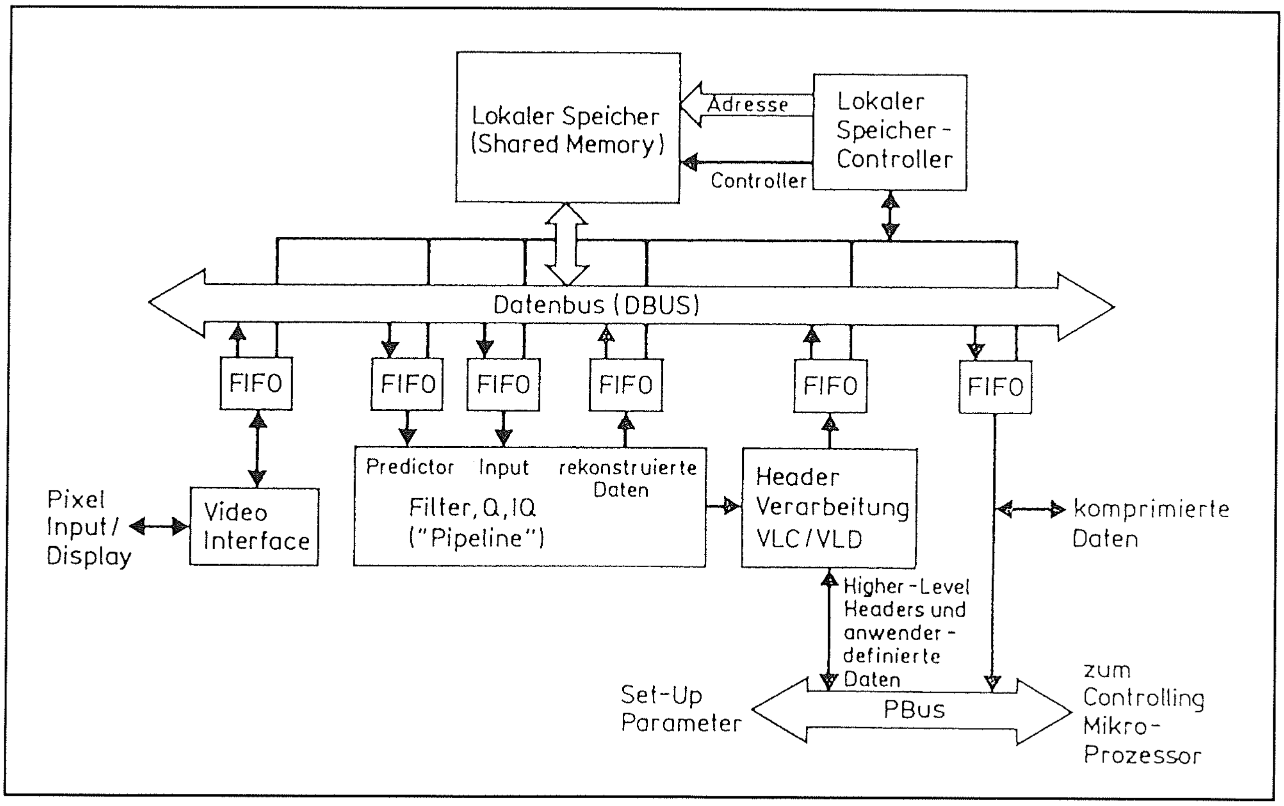


Figuur 5/17.4-4: Voor het verwerken van de hoge resolutie formaten van MPEG-2 moet men overschakelen op parallelle verwerking.

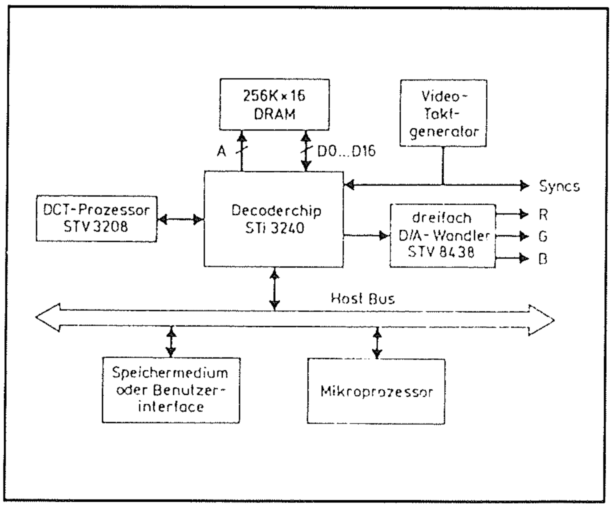
MPEG videodecoders

Een voorbeeld
Als voorbeeld van een decoder voor MPEG-1 wordt de STI3240 van SGC-Thomson blokschematisch besproken. Het blokschema is getekend in figuur 5/17.4-5. De MPEG-decoder moet samenwerken met een externe IDCT-processor, waarvoor bijvoorbeeld de STV3208 kan worden toegepast. Daarnaast is een driekanaals DAC noodzakelijk, waarvoor de STV8438 is ontwikkeld. Het systeemschema is voorgesteld in figuur 5/17.4-6.
De drie basisbestanddelen van de STI3240 zijn de "Compressed Data Unit", de "Decoding Unit" en de "Display Unit". De "Compressed Data Unit" bezit een interface naar de ingangsgegevens, waarmee het mogelijk is de eigenlijke beeldinformatie te scheiden van de gegevens die in de diverse Headers worden verstuurd.

17.4 De MPEG-1 coders en decoders



Figuur 5/17.4-5: Het blokschema van de MPEG-decoder STI3240 van CGS-Thomson.



Figuur 5/17.4-6: Het systeemschema van de STI3240 in combinatie met de IDCT-processor STV3208.

De gegevens worden daarvoor in een register opgeslagen en onderzocht op de aanwezigheid van bepaalde startcodes. De gegevens na de startcodes moeten door de aanwezige systeemsoftware geëvalueerd worden en gebruikt voor het laden van de matrixen voor de dequantisering en de entropiedecodering. De "Decoding Unit" bevat een pipeline geheugen voor de verwerking van de entropiedecodering, de variable lenght decodering en de inverse DCT-functie. De "Dispay Unit" bevat een displaycontroller, die gestuurd moet worden met een externe timinggenerator. De STI3240 kan de videogegevens zowel in RGB- als in YUV-formaat uitvoeren.

Het noodzakelijke externe geheugen met een omvang van 16 keer 256 kB (de in totaal noodzakelijke 4 MB) kan uitgevoerd worden met goedkope standaard DRAM's, die ook gebruikt worden voor het uitbreiden van het extended geheugen van computers.

5/17.5

Video-CD's samenstellen met "DVD PixPlay"

Inleiding

Heeft u ook duizenden foto's?
U heeft waarschijnlijk een digitale fotocamera en een halve harde schijf vol met duizenden foto's in JPG-formaat en filmpjes in AVI-formaat, gemaakt met dezelfde camera. Bekijken van al dat fraais op uw PC is een fluitje van een cent, want zeker de moderne Windows-versie XP heeft prachtige multimedia mogelijkheden ingebouwd. U kunt met de "Windows-viewer voor afbeeldingen en faxen" die automatisch opstart als u op een JPG-bestand in een map dubbelklikt, de volledige set foto's en video's in deze map bekijken. Anders wordt het als u uw foto's ook in de huiskamer op uw TV wilt bewonderen. De JPG's en AVI's naar een CD-R branden heeft weinig zin, want uw DVD-speler ondersteunt zes kansen op de tien niet het afspelen van dergelijke bestanden. Bovendien kunt u er dan geen show van maken met tussengevoegde teksten en flitsende overgangen van de ene naar de andere foto.

Video-CD samenstellen met "DVD PixPlay"
Om er zeker van te zijn dat uw foto's en video's op iedere DVD-speler worden afgespeeld moet u van uw JPG's en AVI's een Video-CD maken. Daar bestaan gespecialiseerde programma's voor die als extra voordeel hebben dat zij toelaten uw foto's in een bepaalde volgorde in een show op te nemen, deze show te voorzien van begeleidende teksten, overgangen tussen uw foto's in te voegen en onder uw foto's muziek te "plakken". Een van deze programma's is "DVD PixPlay" van Xequte Software, een aardig programma dat zeer intuïtief werkt en dat weinig mogelijkheden van het medium Video-CD onbenut laat. In dit hoofdstuk laten wij u kennis maken met dit handig stukje software, dat u snel via Internet en met behulp van uw creditcard goedkoop in huis kunt halen.

Highlights
"DVD PixPlay" heeft de volgende mogelijkheden:

LEES OOK:
Hoofdstuk 5/17.1
Hoofdstuk 5/17.3

- stelt een show samen van uw foto's en video's die u in Video-CD formaat op een CD-R brandt en nadien in iedere DVD-speler kunt afspelen;
- de show kunt u ook schermvullend op uw PC afspelen;
- u kunt uw show begeleiden met muziek of commentaar;
- ondersteunt alle bekende grafische formaten, zoals JPG, JPEG2000, GIF, BMP, PNG en daarnaast ook de "ruwe" bestandsformaten van moderne camera's zoals CRW, CR2 en NEF;
- ondersteunt AVI en MPEG op videogebied;
- ondersteunt MP3 en WAV wat audio betreft;
- via de meegeleverde browser kunt u foto's en video's in de juiste volgorde naar uw show slepen;
- tussen foto's en video's kunt u titelpagina's opnemen;
- heeft meer dan 160 overgangseffecten die u tussen uw foto's kunt invoegen;
- preview-optie op uw PC vóór het branden van de Video-CD;
- ondersteunt DVD, VCD, SVCD, XVCD, XSVCD, MPEG en AVI als uitvoer van de show.

Downloaden en registreren
Het installatiebestand PIXPLAY.EXE bevat versie 2.45, is slechts 3,5 MB groot en kunt u ophalen van www.xequte.com. Het programma werkt onder Windows 95/98/ME/NT/2000/XP. Dit bestand bevat een volledig functionele versie van het programma. De enige beperkingen zijn dat uw show maar vijftig foto's kan bevatten en dat iedere foto wordt ontsierd door een "ongeregistreerde versie"-tekst. Niet erg bruikbaar in de praktijk, maar goed genoeg om u te overtuigen van de capaciteiten van het programma. Het registreren gaat uiteraard per Internet en via creditcard. Voor US$ 29,50 krijgt u via e-mail vrijwel onmiddellijk een serienummer toegestuurd, dat u via de optie "Enter Reg. Key" van het "Help"-menu kunt invoeren. Nadien is het programma gebruiksklaar zonder beperkingen.

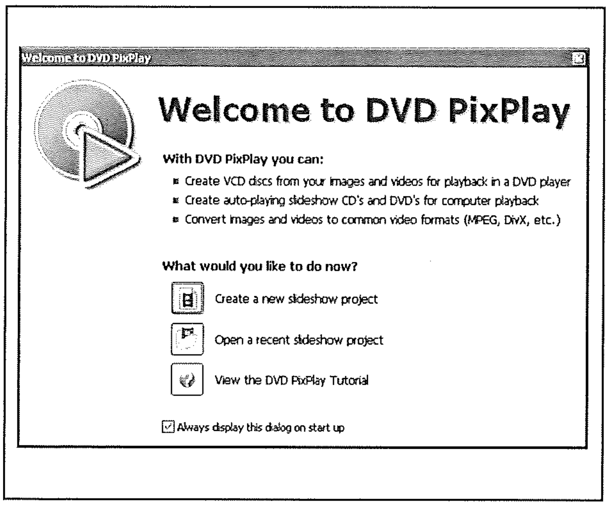
Snel aan de slag!

Voorbereiden
In principe kunt u via "DVD PixPlay" foto's rechtstreeks vanuit de geheugenkaart van uw camera importeren. Wij raden u echter aan de foto's en video's uit de geheugenkaart van uw camera éérst via de bij de camera geleverde software naar een folder van uw harde schijf te kopiëren, dat werkt handiger. Selecteer nadien alle foto's en video's die u op de Video-CD wilt branden en zet deze in een eigen folder. Hernoem de bestanden nadien in de logische volgorde zoals u ze op de Video-CD wilt hebben. De eerste foto van de show hernoemt u dus bijvoorbeeld naar "001.JPG", etc. Dit maakt het in de juiste volgorde opnemen van de foto's in het programma veel gemakkelijker.

De snelle Wizard
Na het opstarten verschijnt de handige Wizard van figuur 5/17.5-1 op uw scherm. Hiermee kunt u op een verbazingwekkend snelle manier een "standaard show" samenstellen. Hiermee bedoelen wij dat deze Wizard u in vijf stappen door een procedure leidt die een bruikbare, maar onpersoonlijke show tot resultaat heeft. Maar na de vijfde stap bestaat gelukkig de mogelijkheid dat u dit snelle resultaat gaat editten, waardoor u

17.5 Video-CD's samenstellen met "DVD PixPlay"

uw show het broodnodige persoonlijke tintje kunt meegeven.
U klikt hier natuurlijk de optie "Create a new slideshow project" aan, waarna u in figuur 5/17.5-2 onmiddellijk de naam van uw show en de locatie op uw harde schijf kunt invullen.



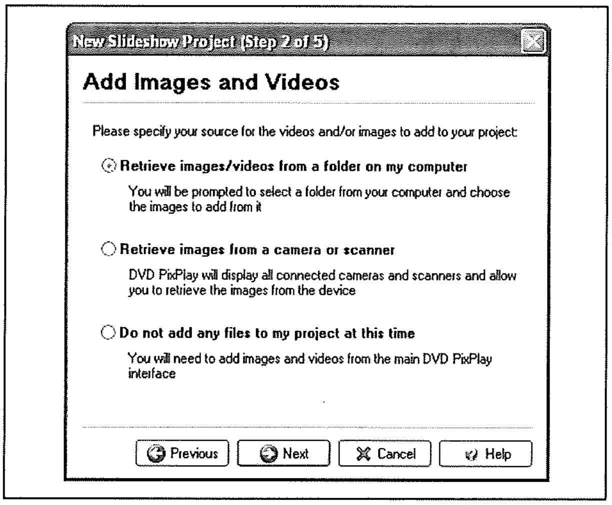
Figuur 5/17.5-1: Aan de hand van deze handige Wizard wordt u in slechts vijf stappen door het samenstellen van een foto- en videoshow geleid.



Figuur 5/17.5-2: Als eerste stap vult u in dit venster de naam van uw project en de locatie ervan op de harde schijf in.

Foto's en video's selecteren

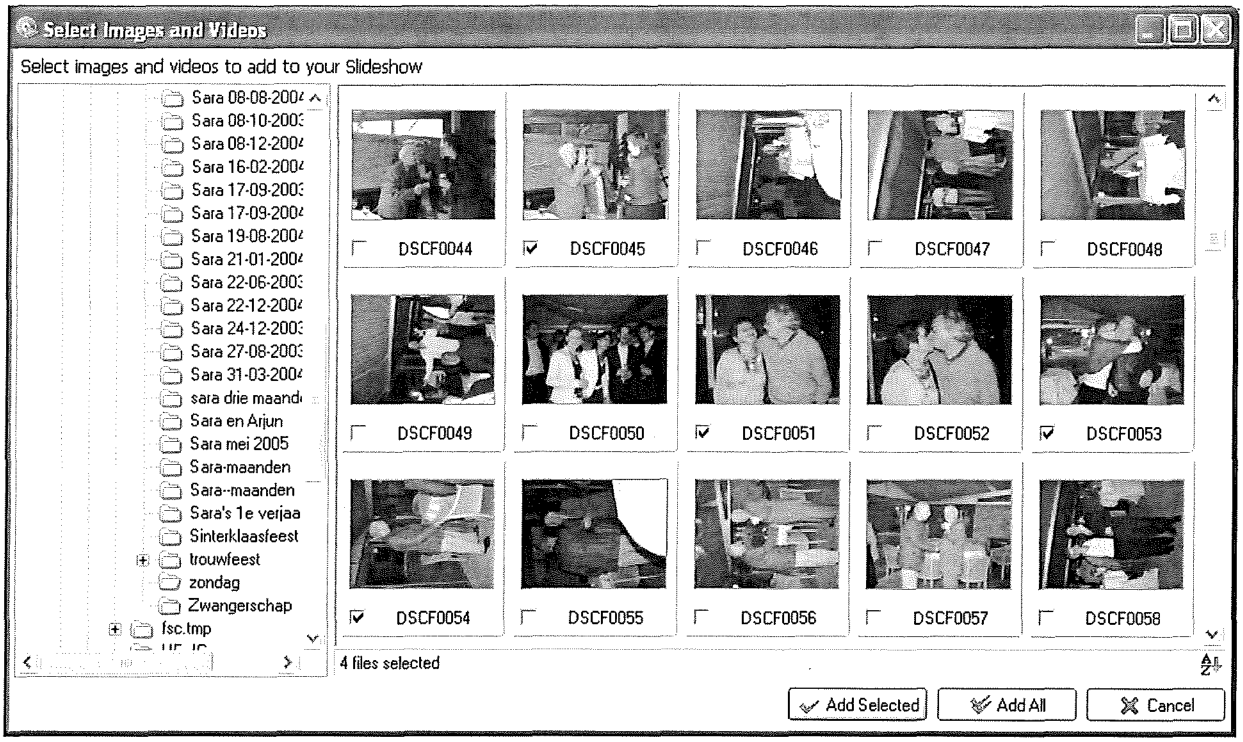
In het volgende venster, voorgesteld in figuur 5/17.5-3, kunt u de bron kiezen waaruit u het materiaal van uw show wilt putten. U kunt kiezen uit een locatie van uw harde schijf ("Retrieve images/videos from a folder on my computer") of rechtstreeks van een camera of scanner ("Retrieve images from a camera or scanner"). Als u onze voorbereidingen heeft opgevolgd is de keuze logisch: optie 1.



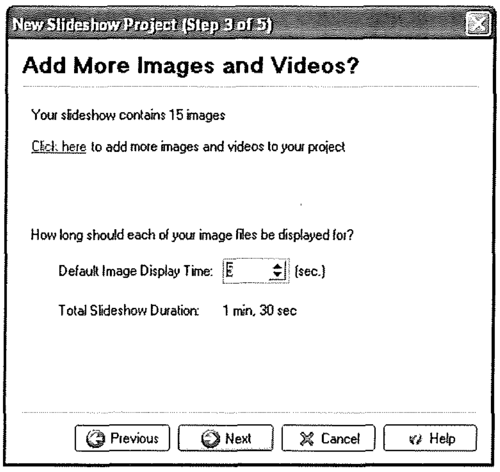
Figuur 5/17.5-3: Het selecteren van de bron van uw materiaal.

Meteen verschijnt het venster "Select Images and Videos" van figuur 5/17.5-4 in beeld, waarin u op de standaard Windows manier foto's uit één (sub)directory kunt aanvinken. Met de knop "Add Selected" voegt u deze foto's aan uw show toe.
In het volgende venster, voorgesteld in figuur 5/17.5-5, krijgt u een overzicht van het aantal geselecteerde foto's. Met de optie "Click here to add more images and videos to your project" kunt u een stapje terug doen en in het venster van figuur 5/17.5-4 een andere directory selecteren waaruit u foto's en video's aan uw show wilt toevoegen.

17.5 Video-CD's samenstellen met "DVD PixPlay"



Figuur 5/17.5-4: Het selecteren van de gewenste fotds uit één Windows-directory.



Figuur 5/17.5-5: Via dit venster kunt u materiaal uit andere mappen toevoegen en de duur van uw show grofweg vastleggen.

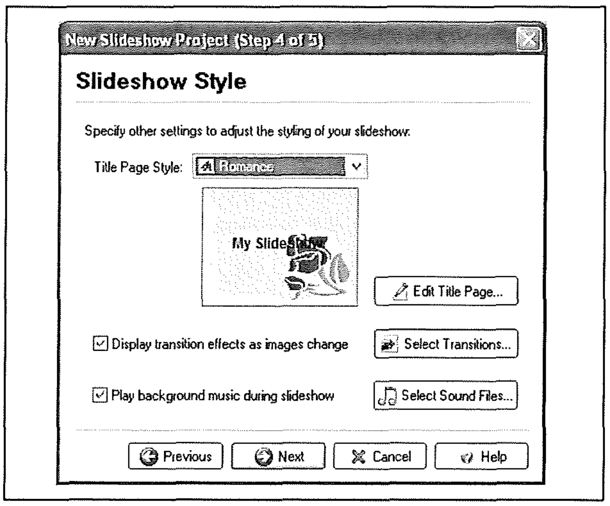
Op deze iets omslachtige manier kunt u dus materiaal uit alle hoeken en gaten van uw hard schijf toevoegen.

Het programma haalt de foto's bij het samenstellen van de Video-CD netjes uit de aangegeven mappen op.
In het venster van figuur 5/17.5-5 kunt u de "Default Image Display Time", oftewel de voorstellingstijd van ieder dia in seconden instellen. U ziet meteen hoe lang uw show gaat duren.

Uw show vormgeven
Nu u het basismateriaal van uw show heeft geselecteerd, kunt u in het volgende venster van de Wizard, de "Slideshow Style", voorgesteld in figuur 5/17.5-6, de broodnodige extra's aanbrengen, waarmee u een saaie opeenvolging van foto's omzet in een échte show. U kiest hier eerst via "Title Page Style" uit een twintigtal ontwerpen voor de titelpagina van uw show. Nadien kunt u met de drie knoppen, rechts in het venster:

17.5 Video-CD's samenstellen met "DVD PixPlay"

- de titelpagina bewerken ("Edit Title Page");
- de overgangen instellen ("Transition Effects");
- eventueel muziek toevoegen ("Select Sound Files").



Figuur 5/17.5-6: In deze voorlaatste stap van de Wizard transformeert u uw saaie verzameling foto's tot een échte show.

Edit Title Page

Na klikken op deze knop verschijnt het venster van figuur 5/17.5-7, waar waarachtig de in de eerste stap ingevoerde titel reeds in de titelpagina is opgenomen. Via dit venster kunt u:
- teksten aan de titelpagina toevoegen, u kunt uiteraard lettertype, -stijl en uitlijning kiezen;
- illustraties toevoegen, het programma bevat een clipart bibliotheek met een honderdtal plaatjes;
- geometrische vormen, van vierkanten tot sterren, aan het ontwerp toevoegen;
- de achtergrond instellen, waarbij graduele overgangen van de ene naar de andere kleur tot de mogelijkheden behoren.



Figuur 5/17.5-7: Het vormgeven van uw titelpagina.

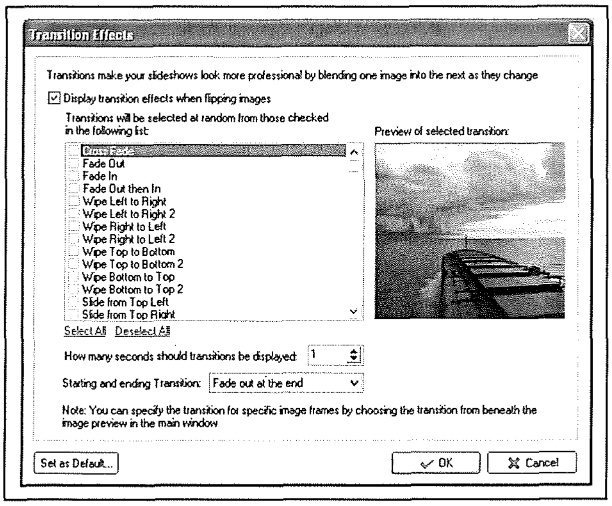
Transition Effects

In dit venster, voorgesteld in figuur 5/17.5-8, kunt u de overgangseffecten invoeren tussen de foto's en video's van uw show. In de lijst staat een honderdtal overgangseffecten ter beschikking, van een simpele "Cross Fade" tot zeer ver gezochte en drukke overgangseffecten. U kunt alle effecten activeren ("Select All") of alleen gewenste effecten aanvinken. Nadien zal het programma de geselecteerde effecten willekeurig tussen de foto's van uw show invoegen. Selecteert u maar één effect, dan wordt dit dus hét effect van uw show. Met "How many seconds should transitions be displayed" kunt u de tijdsduur van de overgangen vastleggen. Met "Starting and Ending Transitions" kunt u de overgangseffecten al dan niet toepassen op de eerste en de laatste dia van de show. In de "Preview of selected transition" ziet u het effect van de geselecteerde overgangen.

Select Sound Files

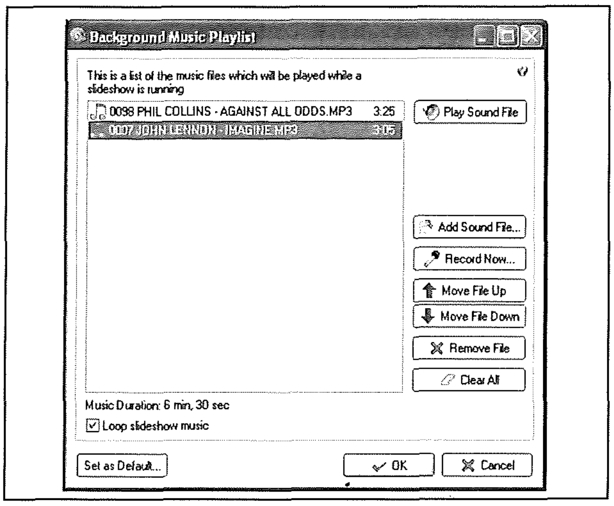
Aan de hand van het venster van figuur 5/17.5-9 kunt u uw show vervolgens begeleiden met muziek.

17.5 Video-CD's samenstellen met "DVD PixPlay"



Figuur 5/17.5-8: In dit venster "Transition Effects" stelt u de overgangen tussen uw foto's en video's in.

Met "Add Sound File" selecteert u de gewenste WAV- of MP3-bestanden op uw harde schijf. Met "Play Sound File" kunt u de geselecteerde bestanden even afspelen.



Figuur 5/17.5-9: Met dit venster "Background Music Playlist" kunt u uw show voorzien van geluid en spraak.

Met "Move File Up" en "Move File Down" kunt u de bestanden in de gewenste volgorde zetten. Met "Loop slideshow music" worden de geluidsbestanden opnieuw afgespeeld als de tijdsduur van uw show langer is dan de totale duur van alle geselecteerde muziekbestanden.
Een interessante optie is "Record Now". Met deze optie kunt u uw show voorzien van commentaar dat u inspreekt via een op uw geluidskaart aangesloten microfoon. Via het venster van figuur 5/17.5-10 kunt u de opname starten. Met "Display Slideshow during the recording" verschijnt de show schermvullend in beeld gedurende het inspreken van het commentaar, zodat u uw commentaar synchroon kunt laten lopen met wat te zien is.



Figuur 5/17.5-10: Via dit venster kunt u uw show van synchroon commentaar voorzien.

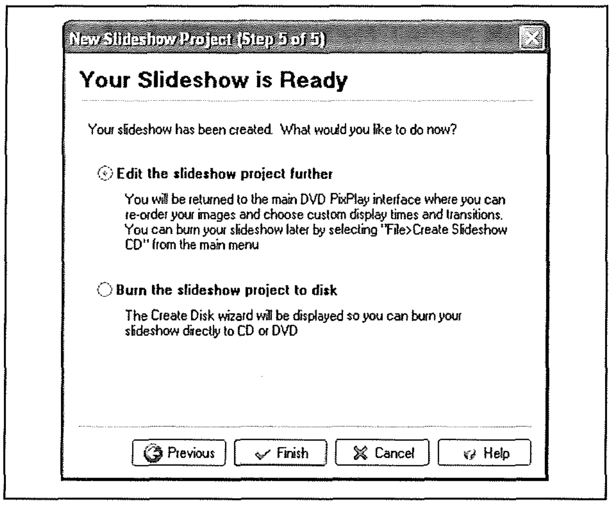
Your Slideshow is Ready

In principe is uw show nu klaar en kan worden omgezet naar een Video-VD. Echter, zoals reeds geschreven, deze procedure levert een standaard show op zonder veel persoonlijke inbreng.
In het vijfde en laatste venster van de Wizard, zie figuur 5/17.5-11, kunt u selecteren uit:

- "Edit the slideshow project further": Met deze optie kunt u uw project gaan "fine tunen".
- "Burn the slideshow to disk":

17.5 Video-CD's samenstellen met "DVD PixPlay"

Uw project wordt omgezet naar Video-CD formaat en op een CD-R gebrand. U doet er verstandig aan voor de eerste optie te kiezen, zodat u uw project kunt verfijnen.



Figuur 5/17.5-11: In dit laatste venster van de Wizard kunt u omschakelen naar het werkvenster van "DVD PixPlay" en uw project verfijnen.

Fine tuning

Het algemeen werkvenster

Kiest u voor de door ons geadviseerde optie, dan verdwijnt het Wizard-venster en komt u in het algemeen werkvenster van figuur 5/17.5-12 terecht. Dit venster biedt alle mogelijke opties voor het personaliseren van uw show. Links in het venster ziet u de geselecteerde foto's onder de vorm van een "Film Strip". Rechts staan de gegevens van de samengestelde show, zoals aantal foto's, duur van de show en geschatte grootte van het totale bestand in Video-VD of DVD-formaat ("Slideshow Details"). Daaronder staan de basisinstellingen die u via de vensters van de Wizard heeft ingevoerd ("Slideshow Setting"). Deze algemene settings kunt u nu echter per foto of video gaan aanpassen. Als u dubbelklikt op één van uw foto's ziet u deze foto als miniatuurweergave rechtsonder in het venster. U kunt nu de "Display Time" en de "Transition" voor de foto individueel aanpassen.

Foto's en video's in de juiste volgorde zetten

Als u onze voorbereidingen heeft opgevolgd staan al uw foto's in de logische volgorde in de show. Maar het kan natuurlijk gebeuren dat u alsnog de volgorde wilt veranderen. Dat is geen probleem. Open het menu "View"en kies in de optie "File Display" voor "Small Tumbs". Alle foto's en video's worden nu geminiaturiseerd weergegeven, zie figuur 5/17.5-13, en u kunt de volgorde van de foto's wijzigen door de foto's naar hun juiste plaats te slepen met ingedrukte muisknop. Verplaatst u een foto naar voren, dan schuiven alle daaropvolgende foto's automatische één plaats op.
Staat al het materiaal definitief in de juiste volgorde, dan kunt u via het menu "View" weer terugschakelen naar "Film Strip" weergave. Dan ziet u weer wat alle foto's voorstellen.

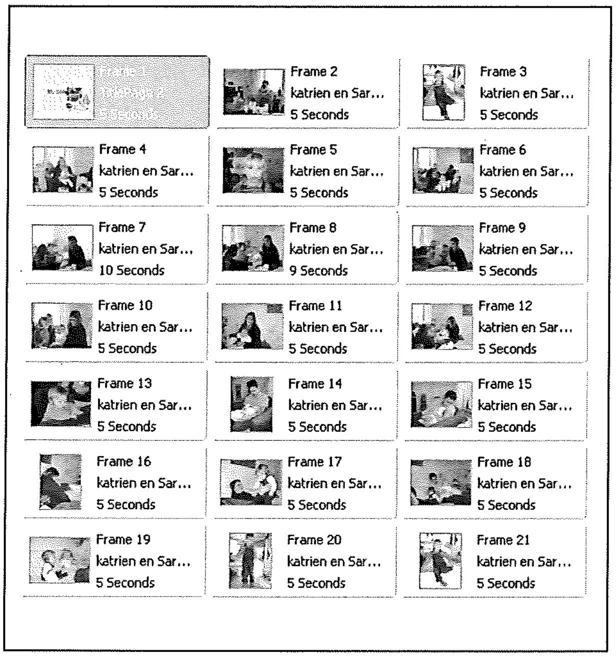
Rechter muisknop

Door met de rechter muisknop op een foto te klikken verschijnt het pop-up venster van figuur 5/17.5-14 in beeld. Een heel handig menu, want u kunt:
- "Add Text to Image":
  Een tekst aan de foto toevoegen.
- "Promote frame":
  De foto één plaats naar voor verplaatsen.
- "Demote Frame":
  De foto één plaats naar achter verplaatsen.

17.5 Video-CD's samenstellen met "DVD PixPlay"



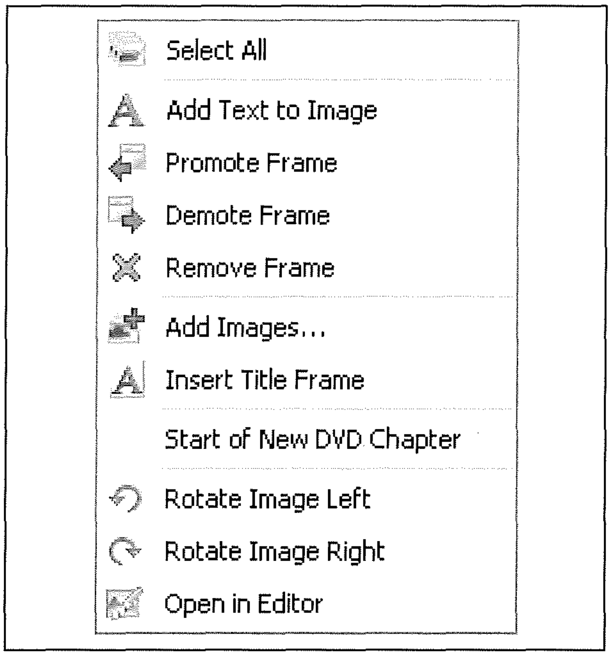
Figuur 5/17.5-12: Het algemeen werkvenster van "DVD PixPlay" waarin u uw show gaat individualiseren.



Figuur 5/17.5-13: In deze miniatuurweergave van uw show kunt u foto's en video's naar een andere plaats verslepen.

- "Remove Frame":
  De foto of video verwijderen.
- "Add Images":
  Extra foto's of video's ná de geselecteerde foto invoegen.
- "Insert Title Frame":
  Een titelpagina ná de geselecteerde foto invoegen;
- "Rotate Image Left":
  De foto 90° naar links draaien.
- "Rotate Image Right":
  Idem, maar dan in de andere richting.
- "Open in Editor":
  De foto openen in de standaard Windows-viewer.
- "Start of New DVD Chapter":
  Als u géén Foto-CD wilt maken maar een échte DVD, dan kunt u aangeven dat deze foto de start is van een nieuw hoofdstuk, waar u via het DVD-menu snel naar toe kunt gaan.

17.5 Video-CD's samenstellen met "DVD PixPlay"



Figuur 5/17.5-14: Dit achter de rechter muisknop verborgen pop-up menu biedt diverse handige functies voor de geselecteerde foto.

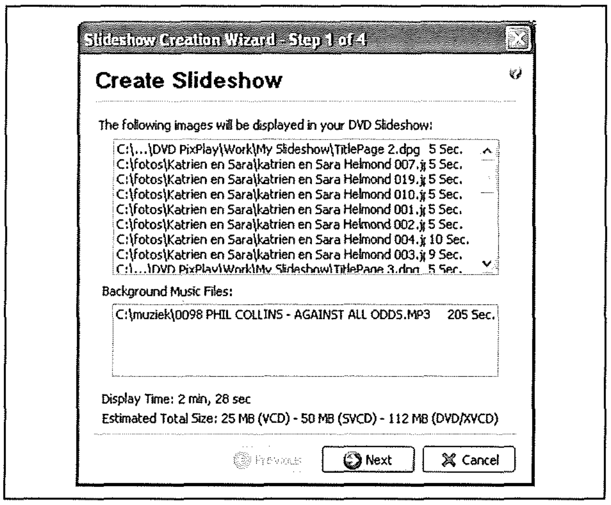
De Video-CD branden

Preview

Als u uw show helemaal heeft samengesteld, kunt u de Video-CD gaan branden. Het is echter aan te raden eerst de "Preview"-knop in de knoppenbalk aan te klikken. Uw show wordt dan beeldvullend afgespeeld en u kunt controleren of alles is zoals het de bedoeling was. Met een druk op "Esc" verlaat u deze preview.

Create Disk

Alles OK? Klik dan op de knop "Create Disk" in de knoppenbalk. In het venster van figuur 5/17.5-15 krijgt u, wellicht ten overvloede, nog eens een overzicht van alle bestanden die u in uw show heeft opgenomen. Dit is de eerste stap van de "Slideshow Creation Wizard", die u in vier stappen begeleidt bij het branden van de Video-CD.



Figuur 5/17.5-15: De eerste stap van de "Slideshow Creation Wizard" geeft nog eens een overzicht van alle geselecteerde bestanden.

Slideshow Format

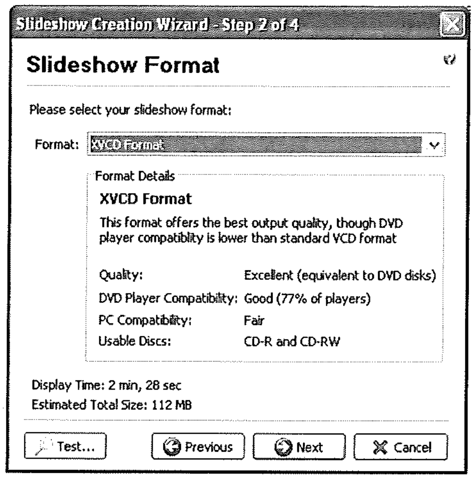
In de volgende stap van de Wizard, voorgesteld in figuur 5/17.5-16, moet u het formaat van de uitvoer specificeren. "DVD PixPlay" biedt zeven uitvoerformaten, ieder met hun eigen voor- en nadelen.

- VCD:
  Dit originele Video-CD formaat heeft als voordeel dat er heel wat foto's op een CD-R passen. Uw show kan maximaal 74 minuten lang zijn. De resolutie van de beelden is echter laag, namelijk 352 bij 288 pixels, te vergelijken met de standaard TV-kwaliteit. Vrijwel iedere DVD-speler ondersteunt dit formaat.
- SVCD:
  Een opgevoerd Video-CD formaat met een resolutie van 480 bij 576 pixels. Er past 35 tot 60 minuten show op één CD-R. Ook dit formaat wordt

17.5 Video-CD's samenstellen met "DVD PixPlay"

door vrijwel alle DVD-speler ondersteund.

- XVCD:
  Video-CD formaat met een resolutie van 720 bij 576 pixels, even hoog als DVD, met als gevolg dat er slechts 15 tot 20 minuten show op één CD-R past. De compatibiliteit met DVD-spelers is echter niet 100 %.

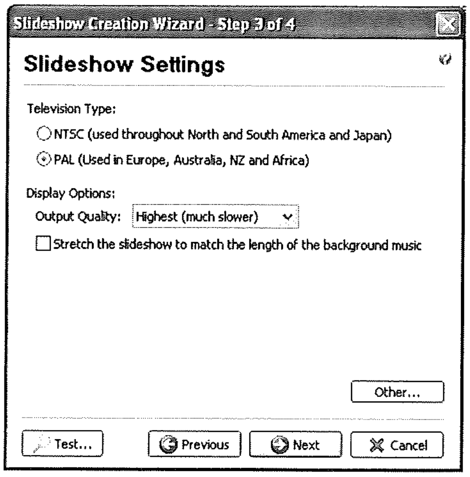


Figuur 5/17.5-16: In stap twee van de "Slideshow Creation Wizard" stelt u het uitvoerformaat in.

- XSVCD:
  Vergelijkbaar met XVCD.
- DVD:
  Heeft dezelfde resolutie als XVCD, maar vanwege het andere medium past er 2 tot 3 uur show op één DVD. U moet natuurlijk wél over een DVD-recorder beschikken.
- Image:
  Schrijft uw originele bestanden naar de CD-R, met als gevolg dat zo'n schijfje waarschijnlijk niet in een DVD-speler is te gebruiken. Dit formaat is echter ideaal voor het verzenden van uw shows naar andere PC-gebruikers die via Windows uw foto's en video's kunnen bewonderen.
- MPEG:
  Het voordeel van dit formaat is dat u een kwaliteitsfactor kunt instellen, waardoor een in feite te lange show toch nog op een CD-R past. Er zijn echter maar weinig DVD-spelers die ruwe MPEG kunnen weergeven. Voor het vertonen op PC's is dit formaat echter ideaal.

Slideshow Setting

In stap drie van deze Wizard stelt u de algemene eigenschappen van de uitvoer in, zie figuur 5/17.5-17. U kiest voor NTSC of PAL, waarbij de tweede keuze uiteraard voor de hand ligt en u kunt nog eens een kwaliteitsfactor instellen.



Figuur 5/17.5-17: In deze derde stap van de output Wizard stelt u een paar gegevens van het uitvoerbestand in.

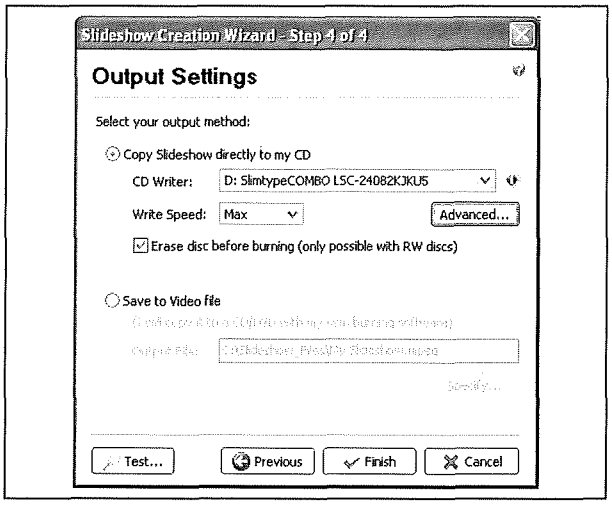
Die keuze heeft de maken met de snelheid waarmee het programma de Video-CD samenstelt. Als u bijvoorbeeld voor VCD kiest, moet het programma alle foto's en video's herschalen naar 352 bij 288 pixels. Het programma beschikt over diverse algoritmen die deze taak uitvoeren. Hoe nauwkeuriger het algorit-

17.5 Video-CD's samenstellen met "DVD PixPlay"

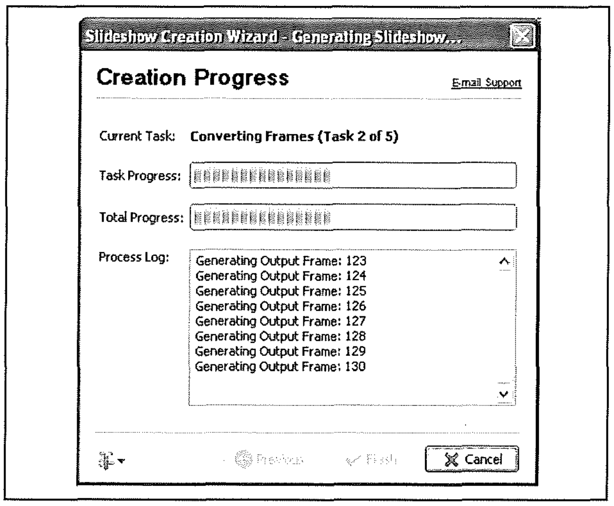
me werkt, hoe trager deze conversie echter gaat. Kiest u voor de beste kwaliteit dan kan, zelfs op een snelle PC, het samenstellen van de Video-CD wel uren duren. Een leuke klus voor uw PC als u op één oor ligt te slapen!
Een handige optie is "Strech the slideshow to match the lenght of the background music". Vinkt u deze optie aan, dan zal het programma de tijdsduur van uw show aanpassen aan de tijdsduur van de door u gekozen achtergrondmuziek. In feite komt het er op neer dat de door u ingestelde "Display Times" iets worden verlengd of verkort.
Met de knop "Test" maakt het programma de outputfile aan, zonder deze echter naar CD-R of DVD-R te schrijven. U kunt dus nogmaals controleren of de show aan uw wensen voldoet.

Output Setting
In de laatste stap van de Wizard, voorgesteld in figuur 5/17.5-18, kunt u kiezen voor rechtstreeks naar een CD-R schrijven ("Copy Slideshow directly to my CD") of voor het schrijven van de Video-CD inhoud naar de harde schijf ("Save to Video file"). In het eerste geval kunt u in het venstertje "CD Writer" uw CD-recorder selecteren. In het laatste geval moet u nadien, via uw eigen brandersoftware, de bestanden naar een CD-R branden.
Met de knop "Advanced" kunt u een aantal opties van het brandproces instellen, maar tenzij u heel goed weet hoe uw CD-recorder werkt, wordt aanbevolen deze opties niet te veranderen.

Creation Process
Het omzetten van uw materiaal naar het video-CD formaat heeft heel wat voeten in de aarde. Op de eerste plaats moeten alle foto's en video's naar de ingestelde resolutie worden geresampled.



Figuur 5/17.5-18: In deze laatste stap kunt u "DVD PixPlay" opdracht geven om rechtstreeks een CD-R te branden.



Figuur 5/17.5-19: Het samenstellen van de Video-CD kunt u in dit venster stap na stap volgen.

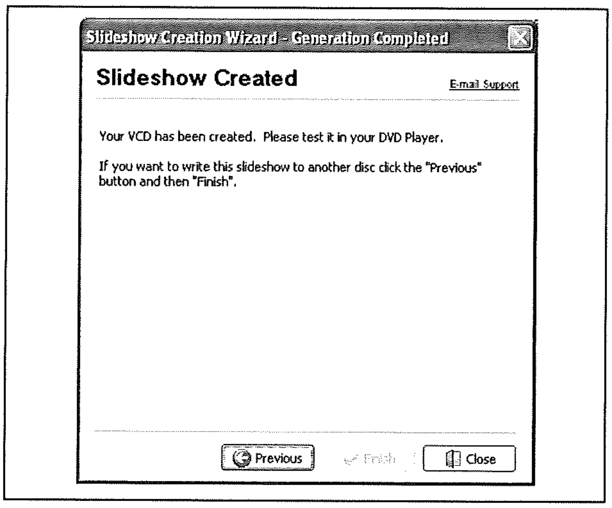
Op de tweede plaats moeten de overgangen tussen de foto's worden berekend en omgezet in grafische bestanden. Nadien moeten al deze gegevens in opeenvolgende frames worden opgenomen. Tot slot moeten deze frames tot

17.5 Video-CD's samenstellen met "DVD PixPlay"

één groot bestand worden verenigd. Aan de hand van het venster dat is voorgesteld in figuur 5/17.5-19 kunt u dit proces volgen.

Slideshow Created

Als u kiest voor de optie "Copy Slideshow directly to my CD" verschijnt, als de CD-R gebrand en afgesloten is, het venster van figuur 5/17.5-20 op uw scherm. Het proces is klaar, de CD-R is omgevormd tot een heuse Video-CD die u op elke DVD-speler kunt vertonen.



Figuur 5/17.5-20: Dit laatste venster meldt dat uw Video-VD productie geslaagd is.